

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/77324 A1

(51) International Patent Classification⁷: **C12N 15/12,**
C07K 14/47, C12N 15/62, A61K 48/00

(74) Agent: **DAVIES, Jonathan, Mark;** Reddie & Grose, 16
Theobalds Road, London WCLX 8PL (GB).

(21) International Application Number: PCT/GB01/01599

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 9 April 2001 (09.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0008582.9 8 April 2000 (08.04.2000) GB

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **ADPRO-TECH LIMITED** [GB/GB]; Second Floor, United 7 & 8, The Maltings, Green Drift, Royston, Herts SG8 5DY (GB).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **STEWART, Michael** [GB/GB]; 4 Meadows Cottages, Thomas Walk, Littlebury Green, Saffron Walden, Essex SG11 4XE (GB). **COX, Vivienne, Frances** [GB/GB]; Bec Cottage, Pounsley Road, Dunton Green, Sevenoaks, Kent TN13 5XP (GB).

(54) Title: DNA IMMUNIZATION VECTORS

(57) Abstract: The invention describes the construction of murine, human and non-human primate variant DNA sequences encoding proteins, such as C3d units, which can be ligated in tandem with each other with or without the native (wild-type) protein (C3d) DNA sequence and may be stably maintained in prokaryotic and eukaryotic expression vectors to produce concatamers of two, three or four copies of either murine, human or non-human primate protein (C3d) at commercially viable levels, and their use in DNA immunization vectors with reduced capacity for homologous integration into host genomic DNA. The invention describes: the construction of novel synthetic DNA sequences encoding concatamers of murine, human or non-human primate C3d where the polypeptide sequence of each unit of the C3d is identical, but the DNA encoding each unit is unique; high-level expression of concatamers of murine, human or non-human primate C3d in prokaryotic and eukaryotic systems and maintenance of stable recombinant expression vector stocks; and the use of variant C3d genes fused to antigen in a DNA immunization vector. The invention also provides a process for preparing oligomeric polypeptides (proteins) in vitro or in vivo which process comprises construction of a DNA vector encoding said polypeptide and its introduction into a recombinant host cell in vitro or host organism in vivo and providing conditions under which said polypeptide will be expressed. The variant DNA polymer comprising a nucleotide sequence that encodes the polypeptide also forms part of the invention.

WO 01/77324 A1

DNA IMMUNIZATION VECTORS

This invention relates to novel genetic constructs designed to permit expression of a naturally-occurring polypeptide from non-native variant DNA sequences, which, when used to express concatamers of the polypeptide show enhanced stability, leading to high level expression in eukaryotic and prokaryotic cell expression systems and when incorporated into a DNA immunization vector reduce the risk of such sequences undergoing homologous recombination with genomic DNA, thus reducing the risk of potentially damaging integration events.

Naturally occurring immune modulators, such as cytokines, or as described below, proteins derived from the complement system, can enhance specific immune responses to an antigen. A number of these have been proposed for inclusion into DNA immunization vectors to be expressed concurrently with the antigen (reviewed by Leitner *et al.*, 1999 Vaccine 18: 765-77). The use of naked DNA as an immunogen has raised concerns about the potential for its integration into the human genome and the possibility of insertional mutagenesis resulting in the inactivation of tumor suppressor genes or the activation of oncogenes (reviewed by Nicholls *et al.*, 1995 Ann N Y Acad Sci 772: 30-9). Although the studies reviewed by Nicholls *et al.*, (1995) have shown this to be a low frequency occurrence with plasmids containing non-human sequences, the inclusion of genes derived from the human genome increases this risk significantly.

This invention may be used in any context where a nucleic acid sequence is included in a medicament where the sequence of the nucleic acid is homologous to a sequence in the genome of the recipient human or animal host. These may be used in the context of gene therapy, therapeutic or prophylactic vaccination or other therapeutic strategies in which nucleic acid forms part of the medicament. It is particularly useful for, but is not restricted to, DNA immunization vectors encoding proteins with immunopotentiating properties derived from the complement system.

The complement system consists of a set of serum proteins that are important in the response of the immune system to foreign antigens. The complement system becomes activated when its primary components are cleaved and the products, alone or with other proteins, activate additional complement proteins resulting in a proteolytic cascade. Activation of the complement system leads to a variety of responses including increased vascular permeability, chemotaxis of phagocytic cells, activation of inflammatory cells, opsonisation of foreign particles, direct killing of cells and tissue damage. Activation of the complement system may be triggered by antigen-antibody complexes (the classical pathway) or a normal slow activation may be amplified in the presence of cell walls of invading organisms such as bacteria and viruses (the alternative pathway). The complement system interacts with the cellular immune system through a specific pathway involving C3, a protein central to both classical and alternative pathways. The proteolytic activation of C3 gives rise to a large fragment (C3b) and exposes a chemically reactive internal thiolester linkage which can react covalently with external nucleophiles such as the cell surface proteins of invading organisms or foreign cells. As a result, the potential antigen is 'tagged' with C3b and remains attached to that protein as it undergoes further proteolysis to iC3b and C3d,g. The latter fragments are, respectively, ligands for the complement receptors CR3 and CR2. Thus the labeling of antigen by C3b can result in a targeting mechanism for cells of the immune system bearing these receptors.

That such targeting is important for augmentation of the immune response is first shown by experiments in which mice were depleted of circulating C3 and then challenged with an antigen (sheep erythrocytes). Removal of C3 reduced the antibody response to this antigen. (M.B. Pepys, J.Exp.Med, 140, 126-145, 1974). The role of C3 was confirmed by studies in animals genetically deficient in either C3 or the upstream components of the complement cascade which generate C3b, i.e. C2 and C4, (J.M. Ahearn & D.T. Fearon, Adv.Immunol. 46, 183-219, 1989). More recently, it has been shown that linear conjugation of a model antigen with more than two copies of the murine C3d fragment sequence resulted in a very large (1000-10000-fold) increase in antibody response in mice compared with unmodified antigen controls (P.W.Dempsey *et al*, Science, 271: 348-350, 1996; WO96/17625, PCT/GB95/02851). The increase could be produced without the use

of conventional adjuvants such as Freund's complete adjuvant. The mechanism of this remarkable effect was demonstrated to be high-affinity binding of the multivalent C3d construct to CR2 on B-cells, followed by co-ligation of CR2 with another B-cell membrane protein, CD19, and with membrane-bound immunoglobulin to generate a signal to the B-cell nucleus.

In the experiments of Dempsey *et al.*, (1996) the unmodified antigen control and linear fusions with one or two C3d domains were prepared by transfection of the appropriate coding plasmids into L cells followed by the selection of high-expressing clones. The most immunogenic construct, that with three C3d units, had to be expressed transiently in COS cells and this procedure gave a very poor yield of the fusion protein. In part, the low yield could be attributed to the generation of species containing the antigen but with lower molecular weights, corresponding to fewer than three C3d units. It was unclear from the published work of Dempsey *et al.* whether the latter molecules originated by proteolysis of the three- C3d construct or whether they were due to a recombination event *in vivo*.

Using another expression system but the same C3d constructs as Dempsey *et al.*, we obtained evidence that the generation of molecules with <3 C3d units from DNA encoding 3x C3d repeats is due to loss of one or more C3d units by homologous recombination and not due to post-translational processing (see WO99/35260) and described methods for the generation and selection of stable variant genes resistant to homologous recombination.

The present invention describes the construction of murine, human and non-human primate variant DNA sequences encoding C3d units which can be ligated in tandem with each other with or without the native (wild-type) C3d DNA sequence and may be stably maintained in prokaryotic and eukaryotic expression vectors to produce concatamers of two, three or four copies of either murine, human or non-human primate C3d at commercially viable levels, and their use in DNA immunization vectors with reduced capacity for homologous integration into host genomic DNA.

The invention comprises the following elements:

1. The construction of novel synthetic DNA sequences encoding concatamers of murine, human or non-human primate C3d where the polypeptide sequence of each unit of the C3d is identical, but the DNA encoding each unit is unique.
2. High-level expression of concatamers of murine, human or non-human primate C3d in prokaryotic and eukaryotic systems and maintenance of stable recombinant expression vector stocks.
3. The use of variant C3d genes fused to antigen in a DNA immunization vector.

The above steps involve the following general processes:

The invention provides a process for preparing oligomeric polypeptides *in vitro* or *in vivo* according to the invention which process comprises construction of a DNA vector encoding said polypeptide and its introduction into a recombinant host cell *in vitro* or host organism *in vivo* and providing conditions under which said polypeptide will be expressed. That process may comprise the steps of:

- (i) preparing a replicable expression vector comprising a nucleotide sequence that encodes said polypeptide;
- ii) transforming a host cell with said vector;
- iii) culturing said transformed host cell under conditions permitting replication of said expression vector or to produce said polypeptide; and
- iv) recovering said expression vector in a form suitable for DNA immunization or said polypeptide in an active form.

The variant genes' DNA polymer comprising a nucleotide sequence that encodes the polypeptide also forms part of the invention.

The process of the invention may be performed using conventional recombinant techniques such as described in Sambrook *et al.*, Molecular Cloning : A laboratory manual 2nd Edition.

Cold Spring Harbor Laboratory Press (1989) and DNA Cloning vols I, II and III (D. M. Glover ed., IRL Press Ltd).

The invention also provides a process for preparing the DNA polymer by the condensation of appropriate mono-, di- or oligomeric nucleotide units.

The preparation may be carried out chemically, enzymatically, or by a combination of the two methods, *in vitro* or *in vivo* as appropriate. Thus, the DNA polymer may be prepared by the enzymatic ligation of appropriate DNA fragments, by conventional methods such as those described by D. M. Roberts *et al.*, in *Biochemistry* 1985, 24, 5090-5098.

The DNA fragments may be obtained by digestion of DNA containing the required sequences of nucleotides with appropriate restriction enzymes, by chemical synthesis, by enzymatic polymerisation, or by a combination of these methods.

Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA. Enzymatic polymerisation of DNA may be carried out *in vitro* using a DNA polymerase such as DNA polymerase 1 (Klenow fragment) in an appropriate buffer containing the nucleoside triphosphates dATP, dCTP, dGTP and dTTP as required at a temperature of 10°-37°C, generally in a volume of 50µl or less. Enzymatic ligation of DNA fragments may be carried out using a DNA ligase such as T4 DNA ligase in an appropriate buffer at a temperature of 4°C to 37°C, generally in a volume of 50µl or less.

The chemical synthesis of the DNA polymer or fragments may be carried out by conventional phosphotriester, phosphite or phosphoramidite chemistry, using solid phase techniques such as those described in 'Chemical and Enzymatic Synthesis of Gene Fragments - A Laboratory Manual' (ed. H.G. Gassen and A. Lang), Verlag Chemie, Weinheim (1982), or in other scientific publications, for example M.J. Gait, H.W.D. Matthes M. Singh, B.S. Sproat and R.C. Titmas, *Nucleic Acids Research*, 1982, 10, 6243; B.S.

Sproat and W. Bannwarth, *Tetrahedron Letters*, 1983, 24, 5771; M.D. Matteucci and M.H. Caruthers, *Tetrahedron Letters*, 1980, 21, 719; M.D. Matteucci and M.H. Caruthers, *Journal of the American Chemical Society*, 1981, 103, 3185; S.P. Adams *et al.*, *Journal of the American Chemical Society*, 1983, 105, 661; N.D. Sinha, J. Biernat, J. McMannus and H. Koester, *Nucleic Acids Research*, 1984, 12, 4539; and H.W.D. Matthes *et al.*, *EMBO Journal*, 1984, 3, 801. Preferably an automated DNA synthesiser (for example, Applied Biosystems 381A Synthesiser) is employed.

The DNA polymer is preferably prepared by ligating two or more DNA molecules which together comprise a DNA sequence encoding the polypeptide. The DNA molecules may be obtained by the digestion with suitable restriction enzymes of vectors carrying the required coding sequences.

The precise structure of the DNA molecules and the way in which they are obtained depends upon the structure of the desired product. The DNA molecule encoding the polypeptide may be constructed using a variety of methods including chemical synthesis of DNA oligonucleotides, enzymatic polymerisation, restriction enzyme digestion and ligation. The design of a suitable strategy for the construction of the DNA molecule coding for the polypeptide is a routine matter for the skilled worker in the art.

The expression of the polypeptide encoded by the DNA polymer in a recombinant host cell or *in vivo* by a recipient of a DNA immunisation vector may be carried out by means of a replicable expression vector capable, in the host cell, of expressing the polypeptide from the DNA polymer.

The replicable expression vector may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or more DNA molecules which, together with said linear segment, encode the polypeptide, under ligating conditions.

The ligation of the linear segment and more than one DNA molecule may be carried out simultaneously or sequentially as desired. Thus, the DNA polymer may be preformed or formed during the construction of the vector, as desired. The choice of vector will be determined in part by the host cell, which may be prokaryotic, such as *E. coli*, mammalian, such as mouse C127, mouse myeloma, Chinese hamster ovary, or other eukaryotic (fungi e.g. filamentous fungi or unicellular yeast or an insect cell such as *Drosophila* or *Spodoptera*). The host cell may also be in a transgenic animal or a human or animal recipient of a DNA immunization vector. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses derived from, for example, baculoviruses, vaccinia, adenovirus and herpesvirus.

The DNA polymer may be assembled into vectors designed for isolation of stable transformed mammalian cell lines expressing the fragment e.g. bovine papillomavirus vectors in mouse C127 cells, or amplified vectors in Chinese hamster ovary cells (DNA Cloning Vol. II D.M. Glover ed. IRL Press 1985; Kaufman, R.J. *et al.* Molecular and Cellular Biology 5, 1750-1759, 1985; Pavlakis G.N. and Hamer, D.H. Proceedings of the National Academy of Sciences (USA) 80, 397-401, 1983; Goeddel, D.V. *et al.* European Patent Application No. 0093619, 1983).

The preparation of the replicable expression vector may be carried out conventionally with appropriate enzymes for restriction, polymerisation and ligation of the DNA, by procedures described in, for example, Sambrook *et al.*, cited above. Polymerisation and ligation may be performed as described above for the preparation of the DNA polymer. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are described in, for example, Sambrook *et al.*, cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985.

The choice of transforming conditions is determined by the host cell. Thus, a bacterial host such as *E. coli*, may be treated with a solution of CaCl_2 (Cohen *et al.*, Proc. Nat. Acad. Sci., 1973, 69, 2110) or with a solution comprising a mixture of RbCl , MnCl_2 , potassium acetate and glycerol, and then with 3-[N-morpholino]-propane-sulphonic acid, RbCl and glycerol or by electroporation as for example described by Bio-Rad Laboratories, Richmond, California, USA, manufacturers of an electroporator. Eukaryotic cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells or by using cationic liposomes. DNA immunization vectors may be administered as naked DNA or contained within a viral particle by injection or by other means of delivery including aqueous or non-aqueous formulations via transdermal or mucosal routes.

The invention also extends to a host cell transformed with a variable replicable expression vector of the invention.

Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Sambrook *et al.*, and "DNA Cloning" cited above. Thus, preferably the cell is supplied with nutrient and cultured at a temperature below 45°C.

The protein product may be recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial such as *E. coli* and the protein is expressed intracellularly, it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. Where the host cell is eukaryotic, the product is usually isolated from the nutrient medium. Where the host cell is in a transgenic animal the protein product may be recovered from the natural secretory pathways (e.g. where the protein is secreted in the milk of a female transgenic animal). Where the host cell is in a human or animal recipient of a DNA immunization vector or gene therapy vector protein products are not normally recovered, but may be detected in tissues for the purpose of evaluating the utility of the delivery system.

WO99/35260 describes methods for purification and refolding (where required) of protein products expressed in prokaryotic and eukaryotic systems.

The nucleic acid may contain an additional cysteine codon which will be expressed at the carboxy-terminus of the polypeptide described in this invention. The utility and post – translational modification of the carboxy-terminal cysteine is described in WO99/35260.

The use of insect cells infected with recombinant baculovirus encoding the polypeptide portion is a preferred general method for preparing complex proteins, particularly the polypeptide encoding C3d oligomers of the invention or fusions of the C3d oligomers with an antigen. The use of DNA immunization vectors is an alternative general method for delivery of the polypeptide encoding C3d oligomers fused to antigen *in vivo* as an immunogen for prophylactic or therapeutic purposes.

GENERAL METHODS USED IN EXAMPLES

(i) DNA Cleavage

Cleavage of DNA by restriction endonucleases was carried out according to the manufacturer's instructions using supplied buffers (New England Biolabs (U.K.) Ltd., Herts. or Promega Ltd., Hants, UK). Double digests were carried out simultaneously if the buffer conditions were suitable for both enzymes. Otherwise double digests were carried out sequentially where the enzyme requiring the lowest salt condition was added first to the digest. Once the digest was complete the salt concentration was altered and the second enzyme added.

(ii) DNA ligation

Ligations were carried out using T4 DNA ligase purchased from Promega or New England Biolabs as described in Sambrook *et al*, (1989) Molecular Cloning: A Laboratory Manual 2nd Edition, Cold Spring Harbor Laboratory Press.

(iii) Plasmid isolation

Plasmids were isolated using Wizard¹⁰™ Plus Minipreps (Promega) or Qiex mini or midi kits and Qiagen Plasmid Maxi kit (QIAGEN, Surrey) according to the manufacturer's instructions.

(iv) DNA fragment isolation

DNA fragments were excised from agarose gels and DNA extracted using the QIAEX gel extraction kit or Qiaquick (QIAGEN, Surrey, UK), or GeneClean, or GeneClean Spin Kit or MERmaid Kit, or MERmaid Spin Kit (Bio 101 Inc, CA. USA) gel extraction kits according to the manufacturer's instructions.

(v) Introduction of DNA into *E. coli*

Plasmids were transformed into competent *E. coli* BL21(DE3) or XL1-blue strains (Studier and Moffat, (1986), J. Mol. Biol. 189:113). The *E. coli* strains were purchased as a frozen competent cultures from Stratagene (Cambridge, UK).

(vi) DNA sequencing

The sequences were analysed by a Perkin Elmer ABI Prism 373 DNA Sequencer. This is an electrophoretic technique using 36 cm x 0.2mm 4% acrylamide gels, the fluorescently labeled DNA fragments being detected by a charge coupled device camera according to the manufacturer's instructions.

(vii) Production of oligonucleotides and synthetic genes

Oligonucleotides and synthetic genes were purchased from Cruachem, Glasgow, UK or from Sigma-Genosys, Cambridge, UK.

(viii) Generation of baculovirus vectors

Plasmids described in this invention having the prefix pBP (e.g. pBP68-03 described below) are used to generate baculovirus vectors and express the encoded recombinant polypeptides by the following methods (Sections (viii) to (x)).

Purified plasmid DNA was used to generate recombinant baculoviruses using the kit 'The BacPak Baculovirus Expression System' according to the manufacturer's protocols

(Clontech, CA, USA). The insect cell line Sf9 (ATCC) was grown in IPL-41 medium (Sigma, Dorset, UK) supplemented according to manufacturers recommendations with yeast extract, lipids and pluronic F68 (all from Sigma) and 1% (v/v) foetal calf serum (Gibco, Paisley, UK) - this is termed growth medium. Cells were transfected with the linearised baculovirus DNA (supplied in the kit) and the purified plasmid. Plaque assays (see method below) were carried out on culture supernatants and a series of ten-fold dilutions thereof to allow isolation of single plaques. Plaques were picked using glass Pasteur pipettes and transferred into 0.5ml aliquots of growth medium. This is the primary seed stock.

(ix) Plaque assay of baculoviruses

1×10^6 Sf9 cells were seeded as monolayer cultures in 30mm plates and left to attach for at least 30 minutes. The medium was poured off and virus inoculum in 100 μ l growth medium was dripped onto the surface of the monolayer. The plates were incubated for 30 minutes at room temperature, occasionally tilting the plates to prevent the monolayer from drying out. The monolayer was overlaid with a mixture of 1ml growth medium and 3% (w/v) "Seaplaque" agarose (FMC, ME) warmed to 37°C and gently swirled to mix in the inoculum. Once set a liquid overlay of 1ml growth medium was applied. The plates were incubated in a humid environment for 3-5 days.

Visualisation of plaques was achieved by addition to the liquid overlay 1ml phosphate buffered saline (PBS) containing neutral red solution at 0.1% (w/v) from a stock solution of 1% (w/v) (Sigma, Dorset, UK). Plaques were visible as circular regions devoid of stain up to 3mm in diameter.

(x) Scale-up of baculovirus vectors and protein expression

200 μ l of the primary seed stock was used to infect 1×10^6 Sf9 monolayer cell cultures in 30mm plates. The seed stock was dripped onto the monolayer and incubated for 20 minutes at room temperature, and then overlaid with 1ml growth medium. The plates were incubated at 27°C in a humid environment for 3-5 days. The supernatant from these cultures is Passage 1 virus stock. The virus titre was determined by plaque assay and further scale up was achieved by infection of monolayer cultures or suspension cultures at a

12

multiplicity of infection (moi) of 0.1. Virus stocks were passaged a maximum of six times to minimise the emergence of defective virus.

Expression of recombinant proteins was achieved by infection of monolayer or suspension cultures in growth medium with or without foetal calf serum (FCS). Where FCS was omitted cells conditioned to growth in the absence of FCS were used. Virus stocks between passage 1 and 6 were used to infect cultures at a moi of >5 per cell. Typically, infected cultures were harvested 72 hours post infection and recombinant proteins isolated either from the supernatants or the cells.

(xi) Protein Purification

A number of standard chromatographic techniques can be used to isolate the C3d-containing proteins, e.g. such methods as ion-exchange and hydrophobic interaction matrixes chromatography utilising the appropriate buffer systems and gradient to purify the target proteins. The properties of the C3d containing fusion polypeptides will vary depending on the nature of the fusion protein. Examples of methods employed in this invention are described in WO99/35260.

(xii) Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was carried out generally using the Novex system (Novex GmbH, Heidelberg) according to the manufacturer's instructions. Pre-packed gels of Tris/glycine a 4- 20% acrylamide gradient were usually used. Samples for electrophoresis, including protein molecular weight standards (for example LMW Kit, Pharmacia, Sweden or Novex Mark 12, Novex, Germany) were usually diluted in 1% (w/v) SDS - containing buffer (with or without 5% (v/v) 2-mercaptoethanol), and left at room temperature for 5 to 30min before application to the gel.

(xiii) Immunoblotting

(a) Dot blot

Immobilon membranes (Millipore, Middlesex, UK) were activated by immersion in methanol for 20 seconds and then washed in PBS for five minutes. The membrane was

placed into a vacuum manifold Dot Blotter (Bio-Rad Laboratories, Watford, UK). Crude extracts from cells or culture supernatants were transferred onto the membrane by applying a vacuum and washed through with PBS. Without allowing the membrane to dry out, the Dot Blotter was dismantled and the membrane removed.

(b) Western Blotting

Samples of cell extracts and purified proteins were separated on SDS-PAGE as described in Section (xii). The Immobilon membrane was prepared for use as in (a) above. The gel and the membrane were assembled in the Semi-Dry Transfer Cell (Trans-Blot SD, Bio-Rad Laboratories) with the Immobilon membrane towards the anode and the SDS-PAGE gel on the cathode side. Between the cathode and the gel were placed 3 sheets of Whatman 3M filter paper cut to the size of the gel pre-soaked in a solution of 192mM 6-amino-n-caproic acid, 25mM Tris pH 9.4 containing 10% (v/v) methanol. Between the anode and the membrane were placed two sheets of Whatman 3M filter paper cut to the size of the gel and soaked in 0.3M Tris pH 10.4 containing 10% (v/v) methanol next to the anode and on this was laid a further sheet of Whatman 3M filter paper pre-soaked in 25mM Tris pH 10.4 containing 10% (v/v) methanol.

The whole-assembled gel assembly was constructed to ensure the exclusion of air pockets. The proteins were transferred from the SDS-PAGE to the Immobilon membrane by passing 200mA current through the assembly for 30 minutes.

(c) Immunoprobng of Dot Blot and Western Membranes

The membranes were blocked by incubating the membrane for 1h at room temperature in 50ml of 10mM phosphate buffer pH 7.4 containing 150mM NaCl, 0.02% (w/v) Ficoll 400, 0.02% (w/v) polyvinylpyrrolidone and 0.1% (w/v) bovine serum albumin (BSA). The appropriate primary antibody was diluted to its working concentration in antibody diluent, 20mM sodium phosphate buffer pH 7.4 containing 0.3M NaCl, 0.5% (v/v) Tween-80 and 1.0% (w/v) BSA. The membrane was incubated for 2h at room temperature in 50ml of this solution and subsequently washed three times for 2 minutes in washing buffer, 20mM sodium phosphate pH 7.4 containing 0.3M NaCl and 0.5% (v/v) Tween-80. The membrane was then transferred to 50ml of antibody diluent buffer containing a suitable

dilution of the species specific antibody labelled with the appropriate label, e.g. biotin, horse radish peroxidase (HRP), for the development process chosen and incubated for 2h at room temperature. The membrane was then washed in washing buffer as described above. Finally, the blot was developed according to the manufacturer's instructions.

The appropriate dilution of antibody for both the primary and secondary antibodies refers to the dilution that minimises unwanted background noise without affecting detection of the chosen antigen using the development system chosen. This dilution is determined empirically for each antibody.

(xiv) Gene sequences

The sequence of wild-type murine and human C3d are available on public databases under accession number K02782 (mouse) and K02765 (human).

EXAMPLES

Example 1: Construction of pBP66-2-14: A baculovirus expression vector encoding a first variant of murine C3d.

pBP66-2-14 is a baculovirus expression vector containing a single copy of murine C3d in which the last 72 amino acid codons contain 41 silent changes resulting in 19% divergence from the wild-type sequence over this region and the 18 amino acid linker region contains 21 further silent changes and two additional amino acids compared to the 16 amino acid linker region used in pBP68-01 (described in WO99/35260). The divergence between the linker in pBP66-2-14 and that in pBP68-01 in the region of silent changes is 59%. The silent changes in the sequence may be third base changes such as substitution of GGG for GGC to encode glycine, or may be two or three base substitutions such as the substitution of AGC for TCC or TCT to encode serine. The gene encoding murine C3d within pBP66-2-14 represents a first variant of murine C3d, (SEQ ID 1) which is designed to be expressed as a dimer or trimer with further variants of murine C3d containing different silent changes. pBP66-2-14 was constructed in five steps as described below.

i) Construction of pBP66-05

The vector pBP66-05 was constructed from pBP66-01 (Described in WO99/35260) using site-directed mutagenesis to introduce a site for HindIII at position 2218 without changing the amino acid sequence. The purpose of this change was to allow direct cloning of a variant gene fragment encoding the carboxy-terminal portion of the murine C3d gene from position 2218 to position 2303. Mutagenesis upon pBP66-01 with the oligonucleotides SEQ ID 2 and SEQ ID 3 and transformation of *E.coli* XL1-blue cells were carried out using the QuikChange Kit (Stratagene) according to the manufacturers instructions. The clone pBP66-05 was selected from transformants by restriction digest analysis of plasmid DNA with the enzyme HindIII.

ii) Construction of pBP66-54-3

The vector pBP66-54-3 was constructed from pBP66-05 using site-directed mutagenesis to introduce multiple silent changes between the positions 2065 and 2218 without changing the amino acid sequence, thus introducing a “fuzzy” gene patch into the wild-type sequence. Four oligonucleotides, Fuz9, Fuz10, Fuz17 and Fuz22 were used to generate a PCR product as described in Example 3a of WO99/35260. The resultant PCR product would have contained a mixed population of “fuzzy” sequences all encoding the same amino acid sequence (apart from those in which PCR errors had arisen). The PCR product was then used to mutagenise pBP66-05 using the QuikChange Kit (Stratagene) according to the manufacturer’s instructions with one variation, whereby said PCR product was used in place of mutagenic oligonucleotides at a final concentration typically in the range 1 to 100 ng/ml. The clone pBP66-54-3 was selected from transformants by restriction digest analysis of plasmid DNA with the enzyme FokI and the integrity of the sequence was confirmed by DNA sequencing.

iii) Construction of pBS-MF2

pBS-MF2 is a holding vector containing the carboxy terminal region of murine C3d cloned from a "fuzzy" PCR product. The fuzzy PCR fragment was derived from four oligonucleotides Fuz11, Fuz12, Fuz21 and Fuz24 as described in WO99/35260. Fuz11 and Fuz12 were overlapping oligonucleotides encoding the carboxy terminal region of murine C3d, plus the linker peptide (Ser-Gly-Gly-Gly-Gly)₂. Fuz21 and Fuz24 were used to amplify the product of Fuz11 and Fuz12. The four oligonucleotides were used to generate a PCR product as described in example 3a of WO99/35260. The resultant PCR product contained a mixed population of "fuzzy" sequences all encoding the same amino acid sequence (apart from those in which PCR errors had arisen). The PCR product was digested with the enzymes HindIII and EagI, and the 124bp fragment was purified by agarose gel electrophoresis and ligated with the large fragment (2912bp) of pBlueScript II KS+ (Stratagene Europe, The Netherlands) digested with the same enzymes and purified in the same way. The ligated DNAs were transformed into *E coli* XL1 blue and resulting transformants were analysed for the insert by PCR screening using oligonucleotides SED ID 4 and SEQ ID 5. The integrity of the sequence was confirmed by DNA sequencing and pBS-MF2 was selected for the next stage of fuzzy murine C3d construction.

iv) Construction of pBP66-54-3/MF#2

The plasmid pBP66-54-3 was subjected to restriction enzyme digestion with the enzymes HindIII and EagI. Two fragments were generated, one of 280bp and one of 450bp. The 450bp fragment was purified by agarose gel electrophoresis.

The plasmid pBS-MF#2 was subjected to restriction enzyme digestion with the enzymes HindIII and EagI. The 175 bp fragment generated was purified by agarose gel electrophoresis and put into ligation reaction with approximately equimolar amount of the 450bp fragment from pBP66-54-3. The ligation reaction was then used as template for a PCR using oligonucleotides SEQ ID 6 and SEQ ID 7.

The primers were designed to amplify up the 625 bp product of the ligation, introducing a HindIII site at either terminus. The product of the PCR was gel extracted and ligated into

the T-vector pCR2.1 (InVitrogen), the sequence determined and the insert then excised with HindIII. This fragment was introduced into the original 54.3 vector which had been digested with HindIII, thus introducing the variant F region onto the end of the variant A-E region to create the plasmid pBP66-54-3/MF#2.

v) Mutagenesis of pBP66-54-3/MF#2 to create pBP66-2-14

Errors, probably arising through PCR, were identified in pBP66-54-3/MF#2. Where repairs were required, oligonucleotides were designed spanning a region of 52 bases around the error. Additional changes were also included in the repair oligonucleotides in order to introduce further silent changes to increase divergence from the sequence of pBP66-01. One change resulted in the introduction of a BsrI restriction site which was used for diagnostic purposes following mutagenesis. The mutagenic oligonucleotides are shown in SEQ ID 8 and SEQ ID 9 and mutagenesis was carried out using the QuikChange Kit (Stratagene) according to the manufacturers instructions. In addition, a Kpn I site was required after the murine C3d coding and the linker region to facilitate subsequent cloning. This was introduced into pBP66-54-3/MF2 by site-directed mutagenesis. The mutagenic oligonucleotides are shown in SEQ ID 10 and 11 and mutagenesis was carried out using the QuikChange Kit (Stratagene) according to the manufacturer's instructions. After the mutagenesis reactions the resulting transformants were subjected to screening by KpnI or BsrI restriction digest analysis of either plasmid DNA or a PCR product spanning the sites of mutagenesis. The integrity of the sequence was confirmed by DNA sequencing.

Example 2: Construction of pBP66-26-15: A Baculovirus expression vector encoding a second variant of murine C3d.

pBP66-26-15 is a baculovirus expression vector containing a single copy of murine C3d in which the first 219 amino acid codons contain 135 silent changes resulting in 20.5% divergence from the wild-type sequence over this region, plus two additional silent changes in the remaining sequence. All the silent changes in pBP66-26-15 were within codons not altered in the corresponding sequence in pBP66-2-14. The silent changes in the sequence

were mostly third base changes such as substitution of GGG for GGC to encode glycine, but also included two or three base substitutions such as the substitution of AGC for TCC or TCT to encode serine. This represents a second variant of murine C3d (SEQ ID 12), which is designed to be expressed as a dimer or trimer with further variants of murine C3d containing different silent changes. The sequence divergence between the 296 amino acid homologous regions of the first and second variants of murine C3d is 20%.

The vector pBP66-26-15 was constructed from pBP66-01 using site-directed mutagenesis to introduce multiple silent changes between the amino acids Thr(T)₁ and Asn(D)₂₁₉ of the murine C3d sequence without changing the amino acid sequence, thus introducing "fuzzy" gene patches into the wild-type sequence. Sixteen oligonucleotides, in four groups of four were used to generate four PCR products as described in example 3a of WO99/35260. Group A contained Fuz1, Fuz2, Fuz23 and Fuz20, Group B contained Fuz3, Fuz4, Fuz19 and Fuz14, Group C contained Fuz5, Fuz6, Fuz13 and Fuz15 and Group D contained Fuz7, Fuz8, Fuz16 and Fuz18. Each of the resultant PCR products A to D would have contained a mixed population of "fuzzy" sequences all encoding the same amino acid sequence (apart from those in which PCR errors had arisen). The PCR products A to D were then used sequentially to mutagenise pBP66-01 using the QuikChange Kit (Stratagene) according to the manufacturer's instructions with one variation, whereby said PCR products were used in place of mutagenic oligonucleotides at a final concentration typically in the range 1 to 100 ng/ml.

After each mutagenesis reaction the resulting transformants were subjected to screening by restriction digest analysis of either plasmid DNA or a PCR product spanning the site of mutagenesis, where introduction of "fuzzy" sequence was presumed to remove restriction sites present in the sequence of pBP66-01. The integrity of the sequence was confirmed by DNA sequencing. One or more clones was selected at each stage for subsequent mutagenesis by another of the four PCR products, or repair of PCR-generated errors where no clones with correct sequence could be identified.

Where repairs were required, oligonucleotides were designed spanning a region of 20-60 bases around the error. Additional changes were also included in such repair oligonucleotides in order to introduce further silent changes to increase divergence from the sequence of pBP66-01 or to introduce or remove restriction sites for diagnostic or cloning purposes. Mutagenesis was carried out using the QuikChange Kit (Stratagene) according to the manufacturer's instructions. After each mutagenesis reaction the resulting transformants were subjected to screening by restriction digest analysis of either plasmid DNA or a PCR product spanning the site of mutagenesis using restriction enzymes corresponding to the diagnostic sites generated by the mutagenesis. The integrity of the sequence was confirmed by DNA sequencing.

When the murine C3d sequence encoded by pBP66-01 had been subjected to silent mutagenesis by all four PCR products A to D, and repaired as required to correct PCR-generated errors, one further mutagenesis reaction was carried out. At the start of the coding sequence it was necessary to alter the reading frame such that the signal peptide and the coding sequence would be in the same frame. A Kpn I site was also introduced to facilitate subsequent cloning. The mutagenic oligonucleotides are shown in SEQ ID 13 and SEQ ID 14 and mutagenesis was carried out using the QuikChange Kit (Stratagene) according to the manufacturers instructions. After each mutagenesis reaction the resulting transformants were subjected to screening by KpnI restriction digest analysis of either plasmid DNA or a PCR product spanning the site of mutagenesis. The integrity of the sequence was confirmed by DNA sequencing.

Example 3: Construction of pBP67-03 containing a third variant of murine C3d.

pBP67-03 is a baculovirus expression vector containing a two copies of murine C3d, the first of which is a third variant of murine C3d (SEQ ID 15) containing 347 changes relative to the wild-type sequence and the second copy is the wild-type murine C3d sequence. The third variant of murine C3d was designed and synthesised *de novo* with the maximum variation at the DNA level from the first and second variants of murine C3d described in

Examples 1 and 2, but encoding an identical polypeptide. The sequence was designed according to the principles of codon variation described in WO99/35260, which takes into account the avoidance of rare codons and was synthesised by Sigma-Genosys (UK), where it was also cloned into the vector pBP66-01 at a unique BglIII site to provide in frame fusion of the two murine C3d fragments and allow expression of a concatameric polypeptide encoding two murine C3d units.

Example 4: Construction pCR-yellow containing a fourth variant of murine C3d.

The first and second variants of murine C3d contain regions of the DNA sequence which are identical to the wild-type sequence. A fourth variant of murine C3d was constructed as a fusion of approximately one third of the sequence from the first variant with approximately two thirds of the sequence from the second variant to generate a sequence containing all the silent changes introduced into both variants. This was achieved by PCR amplification of the variable region from the plasmids pBP66-2-14 containing the first variant and pBP66-26-15 containing the second variant. The oligonucleotide primers used for PCR amplification of pBP66-2-14 are given in SEQ ID 16 and SEQ ID 17, and for PCR amplification of pBP66-26-15 are given in SEQ ID 18 and SEQ ID 19.

The two PCR products were digested with the restriction enzymes BseRI, which cleaves between the wild-type and variant sequence in both PCR products and were purified by gel electrophoresis. The digested fragments containing the variant sequences were ligated *in vitro*. A PCR reaction was carried out to amplify the full-length variant murine C3d sequence using the oligonucleotide primers SEQ ID 16 and SEQ ID 19, and the PCR product was cloned into the vector pCR2.1 (InVitrogen) by T-cloning according to manufacturer's instructions, and the sequence of the resultant plasmid pCR-yellow was authenticated by sequence analysis. The coding sequence of the fourth variant of murine C3d is given in SEQ ID 20.

Example 5: Ligation of three variants of murine C3d in a single concatamer.

The fourth variant of murine C3d was excised from pCR-yellow using the restriction enzymes BglII and BamHI. The 960 base-pair fragment was purified by gel electrophoresis and cloned into the unique BglII site of pBP67-03, which encodes a concatamer of the third (synthetic) variant of murine C3d and wild type murine C3d. The correct orientation of the fourth variant was determined by PCR screening. The resulting plasmid, pBP68-03 is a baculovirus transfer vector containing three copies of murine C3d expressible as a concatamer, where the sequence of each copy differs by 20 – 35%. The sequence of the region of pBP68-03 encoding the murine C3d concatamer and its signal peptide is given in SEQ ID21.

Example 6: Expression of stable murine C3d oligomers in insect cells using pBP68-03.

Expression of murine C3d oligomers using duplicated wild-type sequence in insect cells using the baculovirus expression system was described in WO99/35260, where it was observed that three copies of murine C3d generated a product corresponding to only a single C3d unit, and that the loss of the other two units was due to homologous recombination at the DNA level resulting in deletion of two of the identical DNA sequences encoding the murine C3d units. In this example the plasmids pBP67-03 and pBP68-03 were used to produce recombinant baculoviruses using the methods described above. High levels of murine C3d dimer (including a carboxy-terminal cysteine) were produced by baculoviruses derived from pBP67-03, and of murine C3d trimer by baculoviruses derived from pBP68-03 and the production of the intact oligomeric product was stable over multiple passages of the recombinant baculovirus stock permitting scale-up to large volumes and commercially viable amounts of protein (5-30 mg/litre of culture).

Example 7: Construction of DNA immunization vectors using variant murine C3d sequences.

A model system for DNA immunization in humans is the mouse, where immune responses to antigens produced may be monitored. These models may also be used to determine the frequency of genomic integration events using methods described by Nicholls *et al.*, (1995 Ann N Y Acad Sci 772: 30-9) and the safety profile of such vectors may be evaluated. DNA encoding two of the murine C3d variants with a single copy of wild-type murine C3d

were cloned into the DNA immunization vector pVAX1 (Invitrogen) in tandem with DNA encoding the antigen *Plasmodium yoelii* MSP1.19.

a) construction of pVAX3: A DNA immunization vector for efficient *in vivo* expression of recombinant proteins.

The vector pVAX1 was modified prior to insertion of the murine C3d sequences. The multiple cloning site was removed by digestion with PmeI restriction enzyme and replaced with a synthetic oligonucleotide linker containing the signal peptide sequence from human tissue plasminogen activator (tPA) to create pVAX2. The linker also included sites for BglII and BamHI restriction enzymes, followed by two stop codons. The sequence of the inserted DNA is given in SEQ ID 22 and 23

The pVAX2 vector was subsequently modified by site-directed mutagenesis to generate the "Kozak" consensus sequence (Kozak, M. 1981 Nucleic Acids Res. 9, 5233-62) at the initiation codon of the tPA leader peptide to make the vector pVAX3. The sequence at this point therefore now reads GCCACCATGG.

b) Construction of pVK68-01: A DNA immunization vector encoding murine C3d₃

Murine C3d₃ gene cassettes were introduced into the pVAX3 vector by digestion of the vector with BglII and BamHI. The murine C3d₃ cassette was removed from the baculovirus expression vector pBAC68-04 by digestion with the same enzymes and ligated into the pVAX3 DNA to generate the vector pVK68-01. (pBAC68-04 contains the same C3d₃ cassette as pBP68-03 described above, but the holding vector was pBAC1 (Novagen) instead of pBacPak (Clontech)). Correctly assembled clones of pVK68-01 in pVAX3 were identified by the retention of both BglII and BamHI sites, which could then be used for the insertion of genes encoding antigen. The sequence of pVK68-01 is given in SEQ ID 24

c) Construction of pVK 96-01 and pVK96-02. DNA immunization vectors encoding a malaria antigen fused at the amino or carboxy terminal of murine C3d₃.

A synthetic gene encoding the carboxy-terminal fragment of the *Plasmodium yoelii* MSP1 gene (hereafter described as PyMSP1.19) was synthesised using seven overlapping oligonucleotides, the sequence of which is given in SEQ ID 25 to SEQ ID 31. The amino acid codons within the DNA sequence of PyMSP1.19 were optimized for mammalian expression. The native sequence contains many "rare" codons which were eliminated without affecting the sequence of the encoded polypeptide. The seven oligonucleotides were annealed *in vitro* and then subjected to a two-step PCR amplification using the following method:

1 pmol of each oligonucleotide M1-M7 was incubated with 200uM dNTPs, 0.5 x Taq ligase buffer, 0.5 x Pfu turbo buffer, 5 U Taq ligase and 5 U Pfu turbo polymerase in a total volume of 50 ul. The reaction was subjected to 15 cycles of PCR. The PCR product was then further amplified by the addition of 25 pmol of two primers designed to the termini of the synthetic to 5 ul of the first step PCR reaction, 200 uM dNTPs and 1 x Pfu turbo polymerase in a total volume 50ul for a further 35 PCR cycles.

The reaction product was gel extracted, and 5 ul were incubated with Taq DNA polymerase for 30 minutes at 72°C in the presence of 1 x buffer, MgCl₂ (2mM) and dNTPs (200uM). This promoted the addition of nontemplated A residues to the 3' termini, allowing T-cloning of the products. The product was cloned into the holding vector pUC57/T (MBI Fermentas) and the correct sequence was determined by DNA sequencing.

The PyMSP1.19 gene was excised from the holding vector with BglII and BamHI and inserted into either the BglII site or BamHI site of pVK80-01 to fuse the antigen sequence in frame at either the amino or carboxy termini of the murine C3d₃ cassette. This generated pVK96-01 (amino-terminal fusion) and pVK96-02 (carboxy-terminal fusion) respectively. The sequence of the coding regions for pVK96-01 and pVK96-02 is given in SEQ ID 32 and 33.

d) Analysis of protein products expressed *in vitro* from DNA immunization vectors

Expression of the protein product from the vectors was assessed by carrying out transient transfection of COS7 cells. Using a transfection reagent such as Effectene (Qiagen), the plasmid DNA was introduced into a 90% confluent monolayer of COS7 cells, and incubation of the cells continued for a further 72 hours. Samples of the supernatant and cell pellet were analysed by Western blot for the presence of C3d or antigen-containing protein. All DNA immunization vectors derived by insertion of murine C3d₃ with or without PyMSP1.19 into pVAX3 produced immunoreactive material of the expected molecular weight which was secreted into the supernatant. Efficient expression and secretion from cultured cells *in vitro* provides a degree of confidence that the recombinant fusion proteins will be expressed *in vivo* after the vectors have been administered to mice as described in Example 8.

Example 8. Immunization of mice with DNA immunization vectors encoding murine C3d fused to antigens.

The recombinant DNA immunization vector encoding the murine C3d oligomer-antigen fusions are used to immunize mice using the following protocol. Immunizations are performed using the BioRAD Helios Gene Gun. The plasmid DNA is precipitated onto gold microcarriers in the presence of spermidine, and the gold coated onto the inside of "gold-coat" tubing. The 12.7 mm [0.5"] lengths of tubing are stored desiccated at 4°C until required for use. A single sample of gold-DNA complex is delivered to the shaved abdomen of mice at 2.758 Mpa [400 psi] of helium pressure. A second immunization was performed six weeks after the initial boost.

Vectors encoding more than a single copy of C3d are demonstrated to have enhanced humoral immune responses to the antigen encoded as a fusion to the C3d concatamer. Vectors in which the C3d sequences are non-identical to wild type C3d show a reduced frequency of integration into the genome in comparison with vectors containing wild-type murine C3d.

It can be inferred from these observations that i) a DNA immunization vector encoding greater than one copy of C3d attached to an antigen shows enhanced humoral immune responses to that antigen, and ii) DNA immunization vectors encoding murine proteins are less likely to integrate into the genome of the host if the DNA encoding the murine protein is non-identical to the gene found within the genome of the host.

Example 9: Cloning of wild-type human C3d from a liver library and its expression in *E. coli*.

Human liver total RNA was obtained from Origene Inc. 2 micrograms were used in a conventional reverse transcriptase (RT) reaction with 20pmol oligonucleotide primer SEQ ID 34. The RT reaction was then subjected to 35 cycles of PCR by appropriate adjustment of buffer conditions and addition of the oligonucleotide primer SEQ ID 35. The generation of an amplicon of 892bp as visualized on agarose gel electrophoresis was evidence of successful amplification of the human C3d mRNA. The PCR product was cloned directly into the vector pCR2.1 (InVitrogen) by T-cloning according to manufacturer's instructions, and the sequence of pCR-hC3d-2 authenticated by sequence analysis.

This DNA construct was used as template for a subsequent round of PCR amplification to generate a DNA which possessed additional features to facilitate its subcloning into the *E. coli* expression vector pET26b (Novagen). The following features were introduced at the 5' end of the gene:

- i) The addition of two G bases at the 5' end of the primer to increase the efficiency of T cloning (GG).
- ii) The incorporation of an NdeI site incorporating a synthetic initiation ATG codon. (CATATG).
- iii) The alteration of the CCC codon for Pro (P)₃ to CCG (where ATG is codon 1). CC is a rare codon in *E. coli* and its presence at the start of the gene may have compromised expression levels.
- iv) The alteration of codon 6 from TGC (Cys) to TCC (Ser) to remove the natural site for thioester linkage of C3d to antigenic surfaces.

Features i) to iv) were introduced with the oligonucleotide primer SEQ ID 36.

The following alterations were made at the 3' end of the sequence:

v) The addition of two G bases at the 5' end of the primer to increase the efficiency of T cloning (GG).

vi) The addition of codons for five extra amino acids (SSGSC) and a termination codon (TCAGCAGGATCCACTGCT)

vii) The incorporation of an EcoRI restriction enzyme site (GAATTC)

The alterations v) to vii) were made using the oligonucleotide primer SEQ ID 37.

The product of a standard PCR reaction performed using these primers and pCR-hC3d-2 as template was subcloned into the pCR2.1 vector and a selected clone sequenced. The sequence of the human C3d clone hereafter referred to as the "wild-type" sequence is given in SEQ ID 38.

The insert from pCR78-01 was excised using digestion with NdeI and EcoRI at the sites incorporated into the PCR primers, gel purified and ligated into pET26b, which had been prepared by digestion with the same restriction enzymes within its multiple cloning site. The two DNAs were ligated together and a recombinant clone selected (pET78-01). The clone was verified by restriction mapping prior to transformation into an appropriate expression strain such as BL21(DE3).

Colonies from the transformation of pET78-01 into the expression strain were grown in LB medium until mid-log phase and the expression of the recombinant protein induced by the addition of IPTG to a final concentration of 1mM, followed by a further three hours growth. At the end of the growth period, the cells were harvested and a proportion lysed in reducing NuPAGE sample buffer (Novex) and the proteins analysed by SDS-PAGE on a 10% NuPAGE gel. In addition, the nature of the protein was further examined by the use of a polyclonal antiserum against human C3d (The Binding Site Ltd) by Western blotting, which showed the production of an immunoreactive species at the expected molecular weight.

Example 10: Synthesis of variant human C3d genes

Two variants of human C3d were designed and synthesised *de novo* with the maximum variation at the DNA level from each other and from the wild-type human C3d described in Example 9 but encoding an identical polypeptide. The sequence was designed according to the principles of codon variation described in WO99/35260, which takes into account the avoidance of rare codons. The two genes were synthesised by Sigma-Genosys (Cambridge, UK) as a single concatamer, which was cloned into the holding vector pUC18 to generate the vector pUC78-10. The sequence of the two variants is given in SEQ ID 39 and SEQ ID 40.

Example 11: Ligation of two variants of C3d and one wild-type C3d in a contiguous concatamer in pBP80-02.**a) Construction of pBP78-01, a baculovirus transfer vector encoding a single wild-type C3d sequence.**

The plasmid pET78-01 was digested with BglII and BamHI to excise the wild-type human C3d sequence. The vector pBP68-01 containing the wild-type murine C3d sequence fused to a signal peptide was digested with the same enzymes to remove the murine sequence. The 960 base-pair fragment from pET78-01 and the 5.5 kilobase-pair band from pBP68-01 were purified by gel electrophoresis and ligated to produce pBP78-01, which contains wild-type human C3d. The correct orientation of the insert was determined by PCR screening and the sequence of each junction was determined by DNA sequencing to ensure that a correct in-frame fusion of the signal peptide and human C3d had occurred.

b) Construction of pBP80-01, a baculovirus transfer vector encoding three non-identical, concatameric human C3d sequences.

The two synthetic variants of human C3d were excised from pUC78-10 using the restriction enzymes BglII and BamHI. The 1920 base-pair fragment was purified by gel electrophoresis and cloned into the unique BglII site of pBP78-01, which encodes the wild type human C3d. The correct orientation of the two variants was determined by PCR screening. The resulting plasmid, pBP80-01 is a baculovirus transfer vector containing three copies of human C3d

expressible as a concatamer, where the sequence of each copy differs by approximately 30%. The sequence of the region of pBP80-01 encoding the human C3d concatamer and the signal peptide is given in SEQ ID 41.

Example 12: Expression of stable human C3d monomer and oligomers in insect cells using pBP78-01 and pBP80-01.

The plasmids pBP78-01 and pBP80-01 were used to produce recombinant baculoviruses using the methods described above. High levels of human C3d monomer (including a carboxy-terminal cysteine) were produced by baculoviruses derived from pBP78-01, and of human C3d trimer by baculoviruses derived from pBP80-01 and the production of the intact oligomeric product was stable over multiple passages of the recombinant baculovirus stock permitting scale-up to large volumes and commercially viable amounts of protein (50-100 mg/litre of culture).

Example 13: Construction of a DNA immunization vectors using variant human C3d sequences.

a) Construction of pVAX80-01: A DNA immunization vector encoding human C3d₃.

Human C3d₃ gene cassettes were introduced into the pVAX3 vector by digestion of the vector with BglII and BamHI. The inserts were removed from pBP80-01 (the baculovirus expression vector for human C3d₃) and by digestion with the same enzymes and ligated into the pVAX3 DNA. Correctly assembled clones for human C3d₃ (pVK80-01) in pVAX3 were identified by the retention of both BglII and BamHI sites, which could then be used for the insertion of genes encoding antigen. The sequence of the coding region from pVK80-01 is given in SEQ ID 42.

b) Construction of pVK 104-01 and pVK104-02. DNA immunization vectors encoding a malaria antigen fused at the amino or carboxy terminal of human C3d₃.

The gene for the carboxy terminal fragment of the MSP1 gene from *Plasmodium falciparum* (hereafter described as PfMSP1.19) was obtained as a gift from Dr Anthony Holder, National Institute for Medical Research (London). PfMSP1.19 was PCR amplified and subcloned into a holding vector using primers which introduced a BglII site at the amino terminus and a BamHI site and a Gly-Gly-Gly-Ser-Gly spacer at the carboxy terminus. The sequence of the PfMSP1.19 insert in pUC105-01 is given in SEQ ID 43. It was excised from pUC105-01 with BglII and BamHI and inserted into either the BglII site or BamHI site of pVK80-01 to fuse the antigen sequence in frame at either the amino or carboxy termini of the human C3d₃ cassette. This generated pVK104-01 (amino-terminal fusion) and pVK104-02 (carboxy-terminal fusion) respectively. The sequence of the coding region from pVK104-01 and pVK104-02 are given in SEQ ID 44 and SEQ ID 45.

c) Construction of pVK104-03 and pVK104-04. DNA immunization vectors encoding a variant malaria antigen fused at the amino or carboxy terminal of human Cd3₃.

A variant of PfMSP1.19 was designed such that two cysteine residues that normally form a disulphide bond were converted to other residues. The aim of these mutations was to create an immunogen better able to elicit a protective response than the native amino acid sequence. The rationale for this approach has been described in WO 00/63245. To achieve this the clone pUC105-01 was subjected to site-directed mutagenesis to generate pUC105-03 in two steps, the first to convert Cysteine 12 to Isoleucine, then Cysteine 28 to Tryptophan. The sequence of the altered PfMSP1.19 mutant is given in SEQ ID 46. It was excised from pUC105-03 with BglII and BamHI and inserted into either the BglII site or BamHI site of pVK80-01 to fuse the antigen sequence in frame at either the amino or carboxy termini of the human C3d₃ cassette. This generated pVK104-03 (amino-terminal fusion) and pVK104-04 (carboxy-terminal fusion) respectively. The sequence of the coding region from pVK104-03 and pVK104-04 are given in SEQ ID 47 and SEQ ID 48.

d) Analysis of protein products expressed in vitro from DNA immunization vectors

Expression of the protein product from the vectors was assessed by carrying out transient transfection of COS7 cells. Using a transfection reagent such as Effectene (Qiagen), the plasmid DNA was introduced into a 90% confluent monolayer of COS7 cells, and incubation of the cells continued for a further 72 hours. Samples of the supernatant and cell pellet were analysed by Western blot for the presence of C3d or antigen-containing protein. All DNA immunization vectors derived by insertion of human C3d₃ with or without the native or mutant PfMSP1.19 into pVAX3 produced immunoreactive material of the expected molecular weight which was secreted into the supernatant. Efficient expression and secretion from cultured cells *in vitro* provides a degree of confidence that the recombinant fusion proteins will be expressed *in vivo* after the vectors have been administered as a vaccine to the recipient as described in example 15.

Example 14: Construction of a third variant of human C3d

The vectors PVK80-01, pVK104-01, pVK104-02, pVK104-03 and pVK104-04 contain two variant human C3d genes and one wild-type copy. In order to minimise the risk of integration into the host genome the wild-type gene is replaced with a third variant gene. The third variant is synthesised from overlapping oligonucleotides which, when annealed and amplified produce a synthetic gene with the sequence given in SEQ ID 49. The third variant may also be cloned in tandem with existing sequences encoding three copies of human C3d to generate recombinant proteins with four copies of C3d, either using an expression system such as the baculovirus expression system or in the context of a DNA immunization vector to make the proteins *in vivo*.

Example 15. Immunization of human and non-human primates with DNA immunization vectors encoding human or primate C3d fused to antigens.

The recombinant DNA immunization vectors encoding the human C3d oligomer-antigen fusions are used to immunize humans or non-human primates. The DNA immunization vectors are delivered by a method suitable for delivery of DNA in a clinical protocol, for example, but not restricted to, intramuscular injection. The pVAX1 vector used in the

construction of the DNA immunization vectors described above is suitable for human use and conforms to current FDA guidelines for DNA immunization vectors.

From the studies described in Example 7 it may be inferred that i) a human DNA immunization vectors encoding greater than one copy of human C3d attached to an antigen show enhance humoral immune responses to that antigen, and ii) DNA immunization vectors encoding human proteins are less likely to integrate into the genome of the host if the DNA encoding the human protein is non-identical to the gene found within the genome of the host.

Where it is required for the DNA immunization vectors to be tested in a non-human primate species such as rhesus macaques it may be preferred to use C3d sequences which are exactly matched to the species to be used, in order to minimise immune responses to the C3d component of the encoded protein. The following example (16) describes the cloning of wild-type C3d from a rhesus macaques and the design and synthesis of three species-matched variant genes to use in primate models of human disease to test the safety and immunogenicity of the equivalent DNA immunization vectors encoding human C3d.

Example 16: Cloning of C3d from rhesus liver tissue using degenerate primers

The sequence of the wild-type C3d sequence from rhesus macaque was obtained by cloning the native sequence from liver using the following method:

a) Primer design.

The degenerate primers used to clone the Rhesus macaque-specific C3d sequences were designed by alignment of existing C3 protein sequences from human, mouse, rat, and guinea pig. Regions of amino acid conservation within and flanking the C3d region, where low codon redundancy was prominent were selected by eye, and oligonucleotides for RT-PCR designed to incorporate redundant bases where necessary. The primers, which may be used to clone any mammalian C3d sequence, were designated FARM 1 to FARM 8. The sequence of FARM 1-8 is given in SEQ ID 50 to SEQ ID 57.

b) Reverse transcription-PCR

Total RNA was purified from rhesus macaque liver by the acid-guanidinium thiocyanate-phenol chloroform extraction technique of Chomczynski and Sacchi (Anal. Biochem 162: 156-159 (1987)). Approximately 3ug of RNA was used in the RT reaction using the reverse transcription system from Promega. Reverse transcription was primed with 40pmol of anti-mRNA sense primer, (i.e.. any of the even-numbered primers).

Nested PCR was used to amplify the rhesus macaque C3d in two halves. An outer PCR with primers FARM 4 and 1 was followed by inner PCR with primers FARM 8 and 3 and in a second reaction an outer PCR with primers FARM 5 and FARM 2 were followed by an inner PCR with primers FARM 3 and 8, thus covering the entire C3d region. PCR conditions were typically 95°C 30 sec, 54°C 30 sec, 72°C 60 sec, x 35 cycles.

c) Subcloning and sequencing of novel C3d clones from rhesus macaque.

The PCR products derived from rhesus macaque liver were subcloned into pUC57/T (MBI Fermentas) and a minimum of three clones covering any region of C3d were fully sequenced on both strands. Sequence contigs were assembled and aligned using the SeqMan module of the DNASTar software package. The amino acid sequence of wild-type rhesus macaque C3d is given in SEQID 58.

Example 17: Construction of a three variants of rhesus macaque C3d

The first, second and third variant of rhesus macaque C3d are synthesised from overlapping oligonucleotides which, when annealed and amplified produce three synthetic genes with the sequence given in SEQ ID 59 (first variant), SEQ ID 60 (second variant) and SEQ ID 61 (third variant). As in example 9, the cysteine at position 5 is altered to serine in all the variants to prevent aberrant inter- and intra- molecular disulphide formation. These are used to construct DNA immunization vectors containing variant genes encoding species-matched C3d which have a greatly reduced risk of integration into the host genome over vectors containing non-variant genes encoding species-matched sequences. Such DNA immunization vectors encoding rhesus macaque C3d₃ are fused to antigens which have been

33

selected to produce immune responses for the study of human diseases such as malaria, HIV, hepatitis B virus.

Claims

1. A variant DNA sequence for use in a DNA vaccine, said DNA encoding a naturally occurring protein which, by virtue of third base redundancy and other variations permissible within an amino acid codon, is non-identical to the naturally occurring DNA sequence encoding that protein.
2. A linear concatamer of variant DNA sequences according to claim 1 encoding murine, or human or non-human primate C3d oligomers, and optionally including not more than one wild-type sequence encoding murine, human or other mammalian C3d.
3. Use of a variant DNA sequence according to claim 1 or 2 in a DNA immunization vector to encode one or more naturally occurring human proteins with immunomodulatory properties, said DNA having a sequence which is non-identical to the naturally occurring DNA sequence which encodes that protein.
4. Use of a variant DNA sequence according to claim 1 or 2 in a DNA immunization vector to encode one or more naturally occurring non-human proteins with immunomodulatory properties, said DNA having a sequence which is non-identical to the naturally occurring DNA sequence which encodes that protein.
5. Use of a concatameric C3d sequence according to Claim 2 in a DNA immunization vector for use in murine or primate immunization models or for human immunization, said concatamer being fused to one or more DNA sequences encoding an antigen.
6. A pharmaceutical composition comprising, including or consisting of a variant DNA sequence according to claim 1 or 2 together with a physiologically acceptable excipient or carrier.

7. A pharmaceutical composition comprising, including or consisting of a vector which includes a variant DNA sequence according to claim 1 or 2, together with a physiologically acceptable excipient or carrier.
8. A method of inducing an immune response to an antigen in a human or animal, the method comprising administering a variant DNA sequence according to claim 1 or 2, or administering a vector which includes a variant DNA sequence according to claim 1 or 2.
9. A pharmaceutical composition comprising, including or consisting of a vector as defined in claim 3, 4 or 5 which includes a concatamer according to claim 2, together with a physiologically acceptable excipient or carrier.
10. A method of inducing an immune response to an antigen in the human or animal, the method comprising administering a concatamer according to claim 2, or administering a vector as defined in claim 3, 4 or 5 which includes a concatamer according to claim 2.
11. A method of introducing a DNA sequence encoding a naturally occurring protein into a human or animal, said method comprising administering a pharmaceutical composition according to claims 6, 7, or 9 to said human or animal.
12. A method according to any of claims 8, 10 or 11 wherein said administration results in a therapeutic effect on the human or animal.
13. Use of a variant according to claim 2, or a vector containing said variant, in a method for the manufacture of a medicament for inducing an immune response to an antigen in a human or animal.
14. A vector which comprises, includes or consists of a variant DNA sequence according to claim 1 or 2.

15. A DNA immunization vector comprising or consisting of one or more variant DNA sequences according to claim 1 or 2, fused to one or more DNA sequences encoding an antigen.
16. A variant DNA sequence which encodes human C3d protein, selected from the group consisting of: SEQ ID Nos. 39-42, 44, 45, 47, 48 and 49.
17. A variant DNA sequence which encodes murine C3d protein, selected from the group consisting of: SEQ ID Nos. 1, 12, 15, 20, 21, 24, 32, 33.
18. A variant DNA sequence which encodes non-human primate C3d protein, selected from the group consisting of: SEQ ID Nos. 59-61.
19. An isolated DNA sequence which comprises, includes or consists of SEQ ID No. 58, which DNA sequence encodes rhesus macaque C3d protein.
20. A DNA sequence which is a homolog of a DNA sequence according to any of claims 16 to 19, which homolog differs in sequence by virtue of the addition, deletion or substitution of one or more nucleotide(s) at one or more point(s) or section(s) of said sequence, and wherein said homolog DNA sequence encodes a C3d protein.
21. A vector according to claim 14 which comprises, includes or consists of
 - (i) [m] variant DNA sequences encoding the protein and
 - (ii) [n] wild-type DNA sequences encoding the protein,wherein: [m] = 1, 2, or 3, and;
[n] = 0 or 1.

22. A vector according to claim 20 wherein $[m] = 1, 2 \text{ or } 3$ and $[n] = 0$.
23. A vector according to claim 20 or 21 wherein not all variant sequences are identical.
24. A vector according to claim 20 wherein $[m]=3$ and $[n]=1$ and not all variant sequences are identical.
25. A vector according to claim 20 having three different variant DNA sequences and one wild-type sequence, in any order.
26. A process for preparing an oligomeric polypeptide *in vitro* or *in vivo*, wherein said oligomeric polypeptide is a protein or protein fragment, which process comprises construction of a DNA vector according to claim 14 encoding said polypeptide and its introduction into a recombinant host cell *in vitro* or host organism *in vivo* and providing conditions under which said polypeptide will be expressed: wherein the process comprises the steps of:
- (i) preparing a replicable expression vector comprising a nucleotide sequence according to claim 1 or 2 that encodes said polypeptide;
 - ii) transforming a host cell with said vector;
 - iii) culturing said transformed host cell under conditions permitting replication of said expression vector or to produce said polypeptide; and
 - iv) recovering said expression vector in a form suitable for DNA immunization or said polypeptide in an active form.

DNA sequences described in Examples.**SEQ ID 1: DNA sequence of first murine C3d variant (Example 1)**

```
1  ACCCCCGCAG GCTCTGGGGA ACAGAACATG ATTGGCATGA CACCAACAGT    50
51  CATTGCGGTA CACTACCTGG ACCAGACCGA ACAGTGGGAG AAGTTCGGCA    100
101 TAGAGAAGAG GCAAGAGGCC CTGGAGCTCA TCAAGAAAGG GTACACCCAG    150
151 CAGCTGGCCT TCAAACAGCC CAGCTCTGCC TATGCTGCCT TCAACAACCG    200
201 GCCCCCAGC ACCTGGCTGA CAGCCTACGT GGTCAAGGTC TTCTCTCTAG    250
251 CTGCCAACCT CATCGCCATC GACTCTCACG TCCTGTGTGG GGCTGTTAAA    300
301 TGGTTGATTC TGGAGAAACA GAAGCCGGAT GGTGTCTTTC AGGAGGATGG    350
351 GCCCGTGATT CACCAAGAAA TGATTGGTGG CTTCCGGAAC GCCAAGGAGG    400
401 CAGATGTGTC ACTCACAGCC TTCGTCTCTA TCGCACTGCA GGAAGCCAGG    450
451 GACATCTGTG AGGGGCAGGT CAATAGCCTT CCTGGGAGCA TCAACAAGGC    500
501 AGGGGAGTAT ATTGAAGCCA GTTACATGAA CCTGCAGAGA CCATACACAG    550
551 TGGCCATTGC TGGGTATGCC CTGGCCCTGA TGAACAAACT GGAGGAACCT    600
601 TACCTCGGCA AGTTTCTGAA CACAGCCAAA GATCGGAACC GCTGGGAGGA    650
651 GCCTGACCAG CAGCTCTACA ACGTGGAAGC CACTTCCTAC GCTCTTCTCG    700
701 CACTGCTTCT CCTGAAGGAT TTCGACTCCG TGCCCCCTGT AGTGCGCTGG    750
751 CTGAACGAAC AACGTTACTA TGGGGGGGGG TATGGATCTA CGCAAGCAAC    800
801 ATTCATGGTA TTTCAAGCCT TAGCTCAGTA TCAGACTGAT GTACCAGACC    850
851 ACAAGGATCT TAATATGGAT GTGTCCTTCC ACCTCCCCTC ATCAGGGTCC    900
901 GGAGGGGGTG GATCAGGGGG CGGAGGTTCC GGTACCAGAT CCTAA      935
```

SEQ ID 2:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (i)

1 GGTGTTCCAA GCTTTGGCCC 20

SEQ ID 3:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (i)

1 GGGCCAAAGC TTGGAACACC 20

SEQ ID 4:

DNA sequence of PCR screen oligonucleotide primer used in Example 1 (iii)

1 CAGGAAACAG CTATGAC 17

SEQ ID 5:

DNA sequence of PCR screen oligonucleotide primer used in Example 1 (iii)

1 GTAAAACGAC GGCCAGT 17

SEQ ID 6:

DNA sequence of PCR oligonucleotide primer used in Example 1 (iv)

1 GGAAGCTTTG GCCCAGTATC AGACTGAT 28

SEQ ID 7:

DNA sequence of PCR oligonucleotide primer used in Example 1 (iv)

1 GGAAGCTTTA TTTAACGTGT TTACGTCGAG 30

SEQ ID 8:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (v)

1 GCTTTGCCCC AGTATCAGAC TGATGTACCA GACCACAAGG ATCTTAATAT GG 52

SEQ ID 9:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (v)

1 CCATATTAAG ATCCTTGTGG TCTGGTACAT CAGTCTGATA CTGGGCCAAA GC 52

SEQ ID 10:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (v)

1 GGTTCGGTA CCTGCTAACC 20

SEQ ID 11:

DNA sequence of mutagenic oligonucleotide primer used in Example 1 (v)

1 GGTTAGCAGG TACCGGAACC 20

SEQ ID 12: DNA sequence of second murine C3d variant (Example 2)

1	ACCCCAGCGG	GCTCCGGAGA	ACAAAACATG	ATTGGAATGA	CGCCTACAGT	50
51	CATTGCGGTC	CACTACCTGG	ACCAGACCGA	ACAGTGGGAG	AAATTCGGAA	100
101	TCGAGAAACG	CCAAGAAGCA	CTGGAGCTGA	TTAAAAAGGG	CTATACGCAG	150
151	CAGCTGGCCT	TCAAACAACC	TTCTTCAGCT	TATGCTGCCT	TTAATAACCG	200
201	TCCTCCTTCT	ACGTGGCTTA	CGGCCTACGT	GGTCAAGGTA	TTTTCACTGG	250
251	CAGCTAACCT	CATTGCGATT	GATAGCCACG	TGTTATGCGG	CGCCGTAA	300
301	TGGTTGATT	TCGAGAAGCA	GAAGCCGGAT	GGAGTTTTTC	AAGAAGACGG	350
351	ACCGGTCATT	CACCAAGAGA	TGATTGGTGG	TTTTTCGCAAC	GCCAAGGAGG	400
401	CAGATGTCTC	ACTGACGGCA	TTCGTGCTCA	TCGCGCTTCA	AGAAGCACGT	450
451	GACATTTGCG	AAGGACAAGT	AAACAGCCTT	CCCGGCTCCA	TTAATAAGGC	500
501	TGGTGAGTAC	ATTGAGGCGT	CATATATGAA	TCTTCAACGT	CCTTATACGG	550
551	TCGCTATCGC	GGGCTACGCC	CTGGCCCTCA	TGAACAAACT	TGAGGAACCA	600
601	TACCTAGGAA	AATTCCTGAA	TACAGCCAAG	GATCGTAATC	GTTGGGAGGA	650
651	GCCTGATCAG	CAGCTCTACA	ACGTAGAGGC	CACATCCTAC	GCCCTCCTGG	700
701	CCCTGCTGCT	GCTGAAAGAC	TTTGACTCTG	TGCCCCCTGT	AGTGCGCTGG	750
751	CTCAATGAGC	AAAGATACTA	CGGAGGCGGC	TATGGCTCCA	CCCAGGCTAC	800
801	CTTCATGGTG	TTCCAAGCTT	TGGCCCAATA	TCAAACAGAT	GTCCCTGACC	850
851	ATAAGGACTT	GAACATGGAT	GTGTCCTTCC	ACCTCCCCAG	CAGTGGATCC	900
901	TGCTAGAGTT	CTGA				935

SEQ ID 13: DNA sequence of mutagenic oligonucleotide primer used in Example 2

1 CACCCGAGCC GGTACCAGAT CCACC 25

SEQ ID 14: DNA sequence of mutagenic oligonucleotide primer used in Example 2

1 GGTGGATCTG GTACCGGCTC GGGTG 25

SEQ ID 15:

DNA sequence of third murine C3d variant (Example 3)

1	CCAGATCTAC	GCCTGCCGGT	AGTGGTGAGC	AAAATATGAT	CGGGATGACC	50
51	CCTACTGTGA	TCGCCGTGCA	CTATTTAGAT	CAAACGGAGC	AATGGGAAAA	100
101	ATTTGGGATT	GAAAAACGTC	AGGAAGCGTT	AGAATTGATT	AAAAAGGGAT	150
151	ATACACAACA	ATTAGCGTTT	AAGCAACCAT	CAAGCGCGTA	TGCCGCGTTT	200
201	AACAACAGAC	CACCATCAAC	ATGGTTAACC	GCGTATGTCG	TGAAAGTGTT	250
251	TAGTTTGGCG	GCGAATTTAA	TTGCTATTGA	TAGTCACGTA	TTATGCGGAG	300
301	CGGTAAAGTG	GCTCATCTTA	GAAAAGCAAA	AACCAGACGG	CGTGTTCCAA	350
351	GAAGACGGAC	CAGTCATCCA	CCAGGAGATG	ATCGGGGGCT	TTAGAAATGC	400
401	GAAAGAAGCG	GACGTAAGCT	TAACCGCCTT	TGTATTGATT	GCCTTACAAG	450
451	AGGCGCGCGA	TATTTGCGAA	GGCCAAGTGA	ACTCTTTGCC	GGGATCGATT	500
501	AATAAAGCGG	GCGAATACAT	CGAGGCATCC	TATATGAATT	TACAACGCCC	550
551	TTATACCGTA	GCGATCGCCG	GATACGCGTT	AGCGTTAATG	AATAAGTTAG	600
601	AAGAGCCATA	TTTGGGGAAA	TTCTTAAATA	CGGCGAAGGA	CCGTAATAGG	650
651	TGGGAAGAAC	CAGATCAACA	ATTGTATAAT	GTCGAAGCGA	CCAGTTATGC	700
701	GTTGTTAGCG	TTATTACTTT	TAAAGGATTT	CGATAGCGTC	CCACCAGTGG	750
751	TCAGATGGTT	AAACGAACAG	CGCTATTATG	GCGGAGGTTA	CGGGAGTACA	800
801	CAAGCGACGT	TTATGGTCTT	TCAGGCGCTC	GCGCAGTACC	AGACGGACGT	850
851	GCCAGATCAC	AAAGACCTCA	ATATGGACGT	CAGTTTTTCAC	TTGCCATCAT	900
901	CCGGGAGCGG	CGGAGGTGGG	AGCGGAGGGG	GCGGTACCTC	CGGATCCTAA	950

SEQ ID 16:

DNA sequence of PCR oligonucleotide primer used in Example 4

1 GTGGCTTCCG GAACGCCAA GGAGG 25

SEQ ID 17:

DNA sequence of PCR oligonucleotide primer used in Example 4

1 GCAGGTACCG GATCCTCCGC CCCCTGATCC 30

SEQ ID 18:

DNA sequence of PCR oligonucleotide primer used in Example 4

1 GCCGGTACCA GATCTACCCC AGCGGGCTCC 30

SEQ ID 19:

DNA sequence of PCR oligonucleotide primer used in Example 4

1 GTAGCCTGGG TGGAGCCATA GC 22

SEQ ID 20:

DNA sequence of fourth murine C3d variant (Example 4)

1	CCAGATCTAC	CCCAGCGGGC	TCCGGAGAAC	AAAACATGAT	TGGAATGACG	50
51	CCTACAGTCA	TTGCGGTCCA	CTACCTGGAC	CAGACCGAAC	AGTGGGAGAA	100
101	ATTCGGAATC	GAGAAACGCC	AAGAAGCACT	GGAGCTGATT	AAAAAGGGCT	150
151	ATACGCAGCA	GCTGGCCTTC	AAACAACCTT	CTTCAGCTTA	TGCTGCCTTT	200
201	AATAACCGTC	CTCCTTCTAC	GTGGCTTACG	GCCTACGTGG	TCAAGGTATT	250
251	TTCACCTGGCA	GCTAACCTCA	TTGCGATTGA	TAGCCACGTG	TTATGCGGCG	300
301	CCGTTAAATG	GTTGATTCTC	GAGAAGCAGA	AGCCGGATGG	AGTTTTTCAA	350
351	GAAGACGGAC	CGGTCATTCA	CCAAGAGATG	ATTGGTGGTT	TTCGCAACGC	400
401	CAAGGAGGCA	GATGTCTCAC	TGACGGCATT	CGTGCTCATC	GCGCTTCAAG	450
451	AAGCACGTGA	CATTTGCGAA	GGACAAGTAA	ACAGCCTTCC	CGGCTCCATT	500
501	AATAAGGCTG	GTGAGTACAT	TGAGGCGTCA	TATATGAATC	TTCAACGTCC	550
551	TTATACGGTC	GCTATCGCGG	GCTACGCCCT	GGCCCTCATG	AACAAACTTG	600
601	AGGAACCATA	CCTAGGAAAA	TTCCTGAATA	CAGCCAAGGA	TCGTAATCGT	650
651	TGGGAGGAGC	CTGATCAGCA	GCTCTACAAC	GTGGAAGCCA	CTTCCTACGC	700
701	TCTTCTCGCA	CTGCTTCTCC	TGAAGGATTT	CGACTCCGTG	CCCCCTGTAG	750
751	TGCGCTGGCT	GAACGAACAA	CGTTACTATG	GGGGGGGGTA	TGGATCTACG	800
801	CAAGCAACAT	TCATGGTATT	TCAAGCCTTA	GCTCAGTATC	AGACTGATGT	850
851	ACCAGACCAC	AAGGATCTTA	ATATGGATGT	GTCCTTCCAC	CTCCCCTCAT	900
901	CAGGGTCCGG	AGGGGGTGGA	TCAGGGGGCG	GAGGATCCGT	ACGCAGCTTC	950

SEQ ID 21:

DNA sequence of coding sequence from pBP68-03 from Example 5

1	ATGGCCCTCT	GGATGCGCCT	CCTGCCCCTG	CTGGCCCTGC	TGGCCCTCTG	50
51	GGCGCCCGCG	CCCACCCGAG	CCGTACCAG	ATCTACCCCA	GCGGGCTCCG	100
101	GAGAACAAAA	CATGATTGGA	ATGACGCCCTA	CAGTCATTGC	GGTCCACTAC	150
151	CTGGACCAGA	CCGAACAGTG	GGAGAAATTC	GGAATCGAGA	AACGCCAAGA	200
201	AGCACTGGAG	CTGATTAAAA	AGGGCTATAC	GCAGCAGCTG	GCCTTCAAAC	250
251	AACCTTCTTC	AGCTTATGCT	GCCTTTAATA	ACCGTCCTCC	TTCTACGTGG	300
301	CTTACGGCCT	ACGTGGTCAA	GGTATTTTCA	CTGGCAGCTA	ACCTCATTCG	350
351	GATTGATAGC	CACGTGTTAT	GCGGCGCCGT	TAAATGGTTG	ATTCTCGAGA	400
401	AGCAGAAGCC	GGATGGAGTT	TTTCAAGAAG	ACGGACCGGT	CATTCACCAA	450
451	GAGATGATTG	GTGGTTTTTCG	CAACGCCAAG	GAGGCAGATG	TCTCACTGAC	500
501	GGCATTTCGTG	CTCATCGCGC	TTCAAGAAGC	ACGTGACATT	TGCGAAGGAC	550
551	AAGTAAACAG	CCTTCCCGGC	TCCATTAATA	AGGCTGGTGA	GTACATTGAG	600
601	GCGTCATATA	TGAATCTTCA	ACGTCCTTAT	ACGGTCGCTA	TCGCGGGCTA	650
651	CGCCCTGGCC	CTCATGAACA	AAC TTGAGGA	ACCATACCTA	GGAAAATTCC	700
701	TGAATACAGC	CAAGGATCGT	AATCGTTGGG	AGGAGCCTGA	TCAGCAGCTC	750
751	TACAACGTGG	AAGCCACTTC	CTACGCTCTT	CTCGCACTGC	TTCTCCTGAA	800
801	GGATTTCGAC	TCCGTGCCCC	CTGTAGTGCG	CTGGCTGAAC	GAACAACGTT	850
851	ACTATGGGGG	GGGGTATGGA	TCTACGCAAG	CAACATTCAT	GGTATTTCAA	900
901	GCCTTAGCTC	AGTATCAGAC	TGATGTACCA	GACCACAAGG	ATCTTAATAT	950
951	GGATGTGTCC	TTCCACCTCC	CCTCATCAGG	GTCCGGAGGG	GGTGGATCAG	1000
1001	GGGGCGGAGG	ATCTACGCCT	GCCGGTAGTG	GTGAGCAAAA	TATGATCGGG	1050
1051	ATGACCCCTA	CTGTGATCGC	CGTGCCTAT	TTAGATCAAA	CGGAGCAATG	1100
1101	GGAAAAATTT	GGGATTGAAA	AACGTCAGGA	AGCGTTAGAA	TTGATTAAAA	1150
1151	AGGGATATAC	ACAACAATTA	GCGTTTAAGC	AACCATCAAG	CGCGTATGCC	1200
1201	GCGTTTAACA	ACAGACCACC	ATCAACATGG	TTAACC GCGT	ATGTCGTGAA	1250
1251	AGTGTTTAGT	TTGGCGGCGA	ATTTAATTGC	TATTGATAGT	CACGTATTAT	1300
1301	GCGGAGCGGT	AAAGTGCTC	ATCTTAGAAA	AGCAAAAACC	AGACGGCGTG	1350
1351	TTCCAAGAAG	ACGGACCAGT	CATCCACCAG	GAGATGATCG	GGGGCTTTAG	1400
1401	AAATGCGAAA	GAAGCGGACG	TAAGCTTAAC	CGCCTTTGTA	TTGATTGCCT	1450
1451	TACAAGAGGC	GCGCGATATT	TGCGAAGGCC	AAGTGAAGTC	TTTGCCGGGA	1500
1501	TCGATTAATA	AAGCGGGCGA	ATACATCGAG	GCATCCTATA	TGAATTTACA	1550
1551	ACGCCCTTAT	ACCGTAGCGA	TCGCCGATA	CGCGTTAGCG	TTAATGAATA	1600
1601	AGTTAGAAGA	GCCATATTTG	GGGAAATTCT	TAAATACGGC	GAAGGACCGT	1650
1651	AATAGGTGGG	AAGAACCAGA	TCAACAATTG	TATAATGTCG	AAGCGACCAG	1700
1701	TTATGCGTTG	TTAGCGTTAT	TACTTTTAAA	GGATTTCGAT	AGCGTCCCAC	1750

1751 CAGTGGTCAG ATGGTTAAAC GAACAGCGCT ATTATGGCGG AGGTTACGGG 1800
 1801 AGTACACAAG CGACGTTTAT GGTCTTTCAG GCGCTCGCGC AGTACCAGAC 1850

1851 GGACGTGCCA GATCACAAAG ACCTCAATAT GGACGTCAGT TTTCAGTTGC 1900
 1901 CATCATCCGG GAGCGGCGGA GGTGGGAGCG GAGGGGGCGG TACCTCCGGA 1950
 1951 TCTACCCCCG CAGGCTCTGG GGAACAGAAC ATGATTGGCA TGACACCAAC 2000
 2001 AGTCATTGCG GTACACTACC TGGACCAGAC CGAACAGTGG GAGAAGTTCC 2050
 2051 GCATAGAGAA GAGGCAAGAG GCCCTGGAGC TCATCAAGAA AGGGTACACC 2100
 2101 CAGCAGCTGG CCTTCAAACA GCCCAGCTCT GCCTATGCTG CCTTCAACAA 2150
 2151 CCGGCCCCCC AGCACCTGGC TGACAGCCTA CGTGGTCAAG GTCTTCTCTC 2200
 2201 TAGCTGCCAA CCTCATCGCC ATCGACTCTC ACGTCCTGTG TGGGGCTGTT 2250
 2251 AAATGGTTGA TTCTGGAGAA ACAGAAGCCG GATGGTGTCT TTCAGGAGGA 2300
 2301 TGGGCCCGTG ATTCACCAAG AAATGATTGG TGGCTTCCGG AACGCCAAGG 2350
 2351 AGGCAGATGT GTCCTCACA GCCTTCGTCC TCATCGCACT GCAGGAAGCC 2400
 2401 AGGGACATCT GTGAGGGGCA GGTCAATAGC CTCCTGGGA GCATCAACAA 2450
 2451 GGCAGGGGAG TATATTGAAG CCAGTTACAT GAACCTGCAG AGACCATACA 2500
 2501 CAGTGGCCAT TGCTGGGTAT GCCCTGGCCC TGATGAACAA ACTGGAGGAA 2550
 2551 CCTTACCTCG GCAAGTTTCT GAACACAGCC AAAGATCGGA ACCGCTGGGA 2600
 2601 GGAGCCTGAC CAGCAGCTCT ACAACGTAGA GGCCACATCC TACGCCCTCC 2650
 2651 TGGCCCTGCT GCTGCTGAAA GACTTTGACT CTGTGCCCCC TGTAGTGC GC 2700
 2701 TGGCTCAATG AGCAAAGATA CTACGGAGGC GGCTATGGCT CCACCCAGGC 2750
 2751 TACCTTCATG GTATTCCAAG CCTTGGCCCA ATATCAAACA GATGTCCCTG 2800
 2801 ACCATAAGGA CTTGAACATG GATGTGTCCT TCCACCTCCC CAGCAGTGGA 2850
 2851 TCCTA 2855

SEQ ID 22:

DNA oligonucleotide encoding TPA signal peptide – top strand (example 7)

1 ATGGATGCAA TGAAGAGAGG GCTCTGCTGT GTGCTGCTGC TGTGTGGAGC 50
 51 AGTCTTCGTT TCCGCTAGAT CTGGGTGATA AGGATCCTAG TAA 93

SEQ ID 23:

DNA oligonucleotide encoding TPA signal peptide – bottom strand (example 7)

1	TTACTAGGAT CCTTATCACC CAGATCTAGC GGAAACGAAG ACTGCTCCAC	50
51	ACAGCAGCAG CACACAGCAG AGCCCTCTCT TCATTGCATC CAT	93

SEQ ID 24: DNA sequence of pVK68-01 (murine C3d in pVax1). Coding sequence in bold, signal peptide in italics (Example 7).

1	GACTCTTCGC	GATGTACGGG	CCAGATATAC	GCGTTGACAT	TGATTATTGA	50
51	CTAGTTATTA	ATAGTAATCA	ATTACGGGGT	CATTAGTTCA	TAGCCCATAT	100
101	ATGGAGTTCC	GCGTTACATA	ACTTACGGTA	AATGGCCCCG	CTGGCTGACC	150
151	GCCCAACGAC	CCCCGCCCAT	TGACGTCAAT	AATGACGTAT	GTTCCCATAG	200
201	TAACGCCAAT	AGGGACTTTC	CATTGACGTC	AATGGGTGGA	CTATTTACGG	250
251	TAAACTGCCC	ACTTGGCAGT	ACATCAAGTG	TATCATATGC	CAAGTACGCC	300
301	CCCTATTGAC	GTCAATGACG	GTAAATGGCC	CGCCTGGCAT	TATGCCCAGT	350
351	ACATGACCTT	ATGGGACTTT	CCTACTTGGC	AGTACATCTA	CGTATTAGTC	400
401	ATCGCTATTA	CCATGGTGAT	GCGGTTTGG	CAGTACATCA	ATGGGCGTGG	450
451	ATAGCGGTTT	GACTCACGGG	GATTTCCAAG	TCTCCACCCC	ATTGACGTCA	500
501	ATGGGAGTTT	GTTTTGGCAC	CAAAATCAAC	GGGACTTTCC	AAAATGTCGT	550
551	AACAACGCCG	CCCCATTGAC	GCAAATGGGC	GGTAGGCGTG	TACGGTGGGA	600
601	GGTCTATATA	AGCAGAGCTC	TCTGGCTAAC	TAGAGAACCC	ACTGCTTACT	650
651	GGCTTATCGA	AATTAATACG	ACTCACTATA	GGGAGACCCA	AGCTGGGCCG	700
701	CCACCATGGA	<i>TGCAATGAAG</i>	<i>AGAGGGCTCT</i>	<i>GCTGTGTGCT</i>	<i>GCTGCTGTGT</i>	750
751	<i>GGAGCAGTCT</i>	<i>TCGTTTCCGC</i>	<i>TAGATCTACC</i>	<i>CCAGCGGGCT</i>	<i>CCGGAGAACA</i>	800
801	AAACATGATT	GGAATGACGC	CTACAGTCAT	TGCGGTCCAC	TACCTGGACC	850
851	AGACCGAACA	GTGGGAGAAA	TTCGGAATCG	AGAAACGCCA	AGAAGCACTG	900
901	GAGCTGATTA	AAAAGGGCTA	TACGCAGCAG	CTGGCCTTCA	AACAACCTTC	950
951	TTCAGCTTAT	GCTGCCTTTA	ATAACCGTCC	TCCTTCTACG	TGGCTTACGG	1000
1001	CCTACGTGGT	CAAGGTATTT	TCACTGGCAG	CTAACCTCAT	TGCGATTGAT	1050
1051	AGCCACGTGT	TATGCGGCGC	CGTTAAATGG	TTGATTCTCG	AGAAGCAGAA	1100
1101	GCCGGATGGA	GTTTTTCAAG	AAGACGGACC	GGTCATTAC	CAAGAGATGA	1150
1151	TTGGTGGTTT	TCGCAACGCC	AAGGAGGCAG	ATGTCTCACT	GACGGCATT	1200
1201	GTGCTCATCG	CGCTTCAAGA	AGCACGTGAC	ATTTGCGAAG	GACAAGTAAA	1250
1251	CAGCCTTCCC	GGCTCCATTA	ATAAGGCTGG	TGAGTACATT	GAGGCGTCAT	1300
1301	ATATGAATCT	TCAACGTCCT	TATACGGTCG	CTATCGCGGG	CTACGCCCTG	1350
1351	GCCCTCATGA	ACAAACTTGA	GGAACCATAC	CTAGGAAAAT	TCCTGAATAC	1400
1401	AGCCAAGGAT	CGTAATCGTT	GGGAGGAGCC	TGATCAGCAG	CTCTACAACG	1450
1451	TGGAAGCCAC	TTCTACGCT	CTTCTCGCAC	TGCTTCTCCT	GAAGGATTT	1500
1501	GACTCCGTGC	CCCCTGTAGT	GCGCTGGCTG	AACGAACAAC	GTTACTATGG	1550
1551	GGGGGGGTAT	GGATCTACGC	AAGCAACATT	CATGGTATTT	CAAGCCTTAG	1600
1601	CTCAGTATCA	GACTGATGTA	CCAGACCACA	AGGATCTTAA	TATGGATGTG	1650
1651	TCCTTCCACC	TCCCCTCATC	AGGGTCCGGA	GGGGGTGGAT	CAGGGGGCGG	1700
1701	AGGATCTACG	CCTGCCGGTA	GTGGTGAGCA	AAATATGATC	GGGATGACCC	1750
1751	CTACTGTGAT	CGCCGTGCAC	TATTTAGATC	AAACGGAGCA	ATGGGAAAAA	1800

1801 TTTGGGATTG AAAAACGTCA GGAAGCGTTA GAATTGATTA AAAAGGGATA 1850

1851	TACACAACAA	TTAGCGTTTA	AGCAACCATC	AAGCGCGTAT	GCCGCGTTTA	1900
1901	ACAACAGACC	ACCATCAACA	TGGTTAACCG	CGTATGTCGT	GAAAGTGTTT	1950
1951	AGTTTGGCGG	CGAATTTAAT	TGCTATTGAT	AGTCACGTAT	TATGCGGAGC	2000
2001	GGTAAAGTGG	CTCATCTTAG	AAAAGCAAAA	ACCAGACGGC	GTGTTCCAAG	2050
2051	AAGACGGACC	AGTCATCCAC	CAGGAGATGA	TCGGGGGCTT	TAGAAATGCC	2100
2101	AAAGAAGCGG	ACGTAAGCTT	AACCGCCTTT	GTATTGATTG	CCTTACAAGA	2150
2151	GGCGCGCGAT	ATTTGCGAAG	GCCAAGTGAA	CTCTTTGCCG	GGATCGATTA	2200
2201	ATAAAGCGGG	CGAATACATC	GAGGCATCCT	ATATGAATTT	ACAACGCCCT	2250
2251	TATACCGTAG	CGATCGCCGG	ATACGCGTTA	GCGTTAATGA	ATAAGTTAGA	2300
2301	AGAGCCATAT	TTGGGGAAAT	TCTTAAATAC	GGCGAAGGAC	CGTAATAGGT	2350
2351	GGGAAGAACC	AGATCAACAA	TTGTATAATG	TCGAAGCGAC	CAGTTATGCG	2400
2401	TTGTTAGCGT	TATTACTTTT	AAAGGATTTT	GATAGCGTCC	CACCAGTGGT	2450
2451	CAGATGGTTA	AACGAACAGC	GCTATTATGG	CGGAGGTTAC	GGGAGTACAC	2500
2501	AAGCGACGTT	TATGGTCTTT	CAGGCGCTCG	CGCAGTACCA	GACGGACGTG	2550
2551	CCAGATCACA	AAGACCTCAA	TATGGACGTC	AGTTTTCACT	TGCCATCATC	2600
2601	CGGGAGCGGC	GGAGGTGGGA	GCGGAGGGGG	CGGTACCTCC	GGATCTACCC	2650
2651	CCGCAGGCTC	TGGGGAACAG	AACATGATTG	GCATGACACC	AACAGTCATT	2700
2701	GCGGTACACT	ACCTGGACCA	GACCGAACAG	TGGGAGAAGT	TCGGCATAGA	2750
2751	GAAGAGGCAA	GAGGCCCTGG	AGCTCATCAA	GAAAGGGTAC	ACCCAGCAGC	2800
2801	TGGCCTTCAA	ACAGCCCAGC	TCTGCCTATG	CTGCCTTCAA	CAACCGGCCC	2850
2851	CCCAGCACCT	GGCTGACAGC	CTACGTGGTC	AAGGTCTTCT	CTCTAGCTGC	2900
2901	CAACCTCATC	GCCATCGACT	CTCACGTCCT	GTGTGGGGCT	GTTAAATGGT	2950
2951	TGATTCTGGA	GAAACAGAAG	CCGGATGGTG	TCTTTCAGGA	GGATGGGCCC	3000
3001	GTGATTCACC	AAGAAATGAT	TGGTGGCTTC	CGGAACGCCA	AGGAGGCAGA	3050
3051	TGTGTCACCT	ACAGCCTTCG	TCCTCATCGC	ACTGCAGGAA	GCCAGGGACA	3100
3101	TCTGTGAGGG	GCAGGTCAAT	AGCCTTCCTG	GGAGCATCAA	CAAGGCAGGG	3150
3151	GAGTATATTG	AAGCCAGTTA	CATGAACCTG	CAGAGACCAT	ACACAGTGGC	3200
3201	CATTGCTGGG	TATGCCCTGG	CCCTGATGAA	CAAACCTGGAG	GAACCTTACC	3250
3251	TCGGCAAGTT	TCTGAACACA	GCCAAAGATC	GGAACCGCTG	GGAGGAGCCT	3300
3301	GACCAGCAGC	TCTACAACGT	AGAGGCCACA	TCCTACGCCC	TCCTGGCCCT	3350
3351	GCTGCTGCTG	AAAGACTTTG	ACTCTGTGCC	CCCTGTAGTG	CGCTGGCTCA	3400
3401	ATGAGCAAAG	ATACTACGGA	GGCGGCTATG	GCTCCACCCA	GGCTACCTTC	3450
3451	ATGGTATTCC	AAGCCTTGGC	CCAATATCAA	ACAGATGTCC	CTGACCATAA	3500
3501	GGACTTGAAC	ATGGATGTGT	CCTTCCACCT	CCCCAGCAGT	GGATCCTAGT	3550
3551	AAAAACCCGC	TGATCAGCCT	CGACTGTGCC	TTCTAGTTGC	CAGCCATCTG	3600
3601	TTGTTTGCCC	CTCCCCCGTG	CCTTCCTTGA	CCCTGGAAGG	TGCCACTCCC	3650
3651	ACTGTCCTTT	CCTAATAAAA	TGAGGAAATT	GCATCGCATT	GTCTGAGTAG	3700

3701	GTGTCATTCT	ATTCTGGGGG	GTGGGGTGGG	GCAGGACAGC	AAGGGGGAGG	3750
3751	ATTGGGAAGA	CAATAGCAGG	CATGCTGGGG	ATGCGGTGGG	CTCTATGGCT	3800
3801	TCTACTGGGC	GGTTTTATGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	3850
3851	GCGCCCTCTG	GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	GGATGGCTTT	3900
3901	CTCGCCGCCA	AGGATCTGAT	GGCGCAGGGG	ATCAAGCTCT	GATCAAGAGA	3950
3951	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	AAGATGGATT	GCACGCAGGT	4000
4001	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	GGGCACAACA	4050
4051	GACAAATCGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	4100
4101	GCCCGGTTCT	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	4150
4151	CAAGACGAGG	CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	4200
4201	CGCAGCTGTG	CTCGACGTTG	TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	4250
4251	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	TGCTCCTGCC	4300
4301	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	ATACGCTTGA	4350
4351	TCCGGCTACC	TGCCCATTCT	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	4400
4401	CACGTACTCG	GATGGAAGCC	GGTCTTGTCT	ATCAGGATGA	TCTGGACGAA	4450
4451	GAGCATCAGG	GGCTCGCGCC	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGAG	4500
4501	CATGCCCCGAC	GGCGAGGATC	TCGTCTGTAC	CCATGGCGAT	GCCTGCTTGC	4550
4551	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	CTGGATTTCAT	CGACTGTGGC	4600
4601	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	CTACCCGTGA	4650
4651	TATTGCTGAA	GAGCTTGCGG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	4700
4701	ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	4750
4751	GACGAGTTCT	TCTGAATTAT	TAACGCTTAC	AATTTCTCTGA	TGCGGTATTT	4800
4801	TCTCCTTACG	CATCTGTGCG	GTATTTTACA	CCGCATACAG	GTGGCACTTT	4850
4851	TCGGGGAAAT	GTGCGCGGAA	CCCCTATTTG	TTTATTTTTC	TAAATACATT	4900
4901	CAAATATGTA	TCCGCTCATG	AGACAATAAC	CCTGATAAAT	GCTTCAATAA	4950
4951	TAGCACGTGC	TAAAACTTCA	TTTTTAATTT	AAAAGGATCT	AGGTGAAGAT	5000
5001	CCTTTTTGAT	AATCTCATGA	CCAAAATCCC	TTAACGTGAG	TTTTCGTTCC	5050
5051	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA	AAGGATCTTC	TTGAGATCCT	5100
5101	TTTTTCTGTC	GCGTAATCTG	CTGCTTGCAA	ACAAAAAAAC	CACCGCTACC	5150
5151	AGCGGTGGTT	TGTTTGCCGG	ATCAAGAGCT	ACCAACTCTT	TTTCCGAAGG	5200
5201	TAACTGGCTT	CAGCAGAGCG	CAGATACCAA	ATACTGTCCT	TCTAGTGTAG	5250
5251	CCGTAGTTAG	GCCACCACTT	CAAGAAGTCT	GTAGCACCGC	CTACATACCT	5300
5301	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC	TGCCAGTGGC	GATAAGTCGT	5350
5351	GTCTTACCGG	GTTGGACTCA	AGACGATAGT	TACCGGATAA	GGCGCAGCGG	5400
5401	TCGGGCTGAA	CGGGGGGTTC	GTGCACACAG	CCCAGCTTGG	AGCGAACGAC	5450
5451	CTACACCGAA	CTGAGATACC	TACAGCGTGA	GCTATGAGAA	AGCGCCACGC	5500
5501	TTCCCGAAGG	GAGAAAGGCG	GACAGGTATC	CGGTAAGCGG	CAGGGTCGGA	5550
5551	ACAGGAGAGC	GCACGAGGGA	GCTTCCAGGG	GGAAACGCCT	GGTATCTTTA	5600

5601 TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTGTGAT 5650
5651 GCTCGTCAGG GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT 5700
5701 TTACGGTTCC TGGGCTTTTG CTGGCCTTTT GCTCACATGT TCTT 5744

SEQ ID 25:

DNA sequence of oligonucleotide M1 used in synthesis of PyMSP1.19.

1 AGATCTCACA TTGCCTCTAT TGCTTTGAAC AACTTGAACA AGTCTGGTTT 50
51 GGTAGGAGAA GGTGAGTCTA AGAAGATTTT 80

SEQ ID 26:

DNA sequence of oligonucleotide M2 used in synthesis of PyMSP1.19.

1 GTTGACCCTA AGCATGTTTG TGTGAGACT AGAGACATTC CTAAGAACGC 50
51 TGGATGTTTC AGAGACGACA ACGGTACTGA 80

SEQ ID 27:

DNA sequence of oligonucleotide M3 used in synthesis of PyMSP1.19.

1 AACACCTGCG TTGAGAACAA CAACCCTACT TGCGACATCA ACAACGGTGG 50
51 ATGTGACCCA ACCGCCTCTT GTCAAAACGC 80

SEQ ID 28:

DNA sequence of oligonucleotide M4 used in synthesis of PyMSP1.19.

1 AAGGAACCAA CCCCTAACGC CTACTACGAG GGTGTTTTCT GTTCTTCTTC 50
51 CGGATCC 57

SEQ ID 29:

DNA sequence of oligonucleotide M5 used in synthesis of PyMSP1.19.

1 GCGTTAGGGG TTGGTTCCTT ACAGGTGCAA ATAATCTTCT TGGAGTTTTT 50
51 GGTAGATTCA GCGTTTTGAC AAGAGGCGGT 80

SEQ ID 30:

DNA sequence of oligonucleotide M6 used in synthesis of PyMSP1.19.

```

1      TTGTTCTCAA CGCAGGTGTT ACCCTCACCC TTCTTGTAAC CCAACAAACA      50
51     TCTCCACTCT TCAGTACCGT TGTCGTCTCT      80

```

SEQ ID 31:

DNA sequence of oligonucleotide M7 used in synthesis of PyMSP1.19.

```

1      CAAACATGCT TAGGGTCAAC ACCCAACAAG TCCATACCGT CCATGTTTCA      50
51     CATCTTAGCC AAAATCTTCT TAGACTCACC      80

```

SEQ ID 32: Coding sequence from pVK96-01 (PyMSP1.19-murine C3d₃ in PVAX3 (antigen sequence underlined) from example 7).

```

706     ....ATGGA TGCAATGAAG AGAGGGCTCT GCTGTGTGCT GCTGCTGTGT      750
751     GGAGCAGTCT TCGTTTCCGC TAGATCTCAC ATTGCCTCTA TTGCTTTGAA      800
801     CAACTTGAAC AAGTCTGGTT TGGTAGGAGA AGGTGAGTCT AAGAAGATTT      850
851     TGGCTAAGAT GCTGAACATG GACGGTATGG ACTTGTTGGG TGTTGACCCT      900
901     AAGCATGTTT GTGTTGACAC TAGAGACATT CCTAAGAACG CTGGATGTTT      950
951     CAGAGACGAC AACGGTACTG AAGAGTGGAG ATGTTTGTG GGTTACAAGA     1000
1001    AGGGTGAGGG TAACACCTGC GTTGAGAACA ACAACCCTAC TTGCGACATC     1050
1051    AACACCGGTG GATGTGACCC AACCGCCTCT TGTCAAAACG CTGAATCTAC     1100
1101    CGAAAACCTC AAGAAGATTA TTTGCACCTG TAAGGAACCA ACCCCTAACG     1150
1151    CCTACTACGA GGGTGTTTTC TGTCTTCTT CCGGATCTAC CCCAGCGGGC     1200
1201    TCCGGAGAAC AAAACATGAT TGGAATGACG CCTACAGTCA TTGCGGTCCA     1250
1251    CTACCTGGAC CAGACCGAAC AGTGGGAGAA ATTCGGAATC GAGAAACGCC     1300
1301    AAGAAGCACT GGAGCTGATT AAAAAGGGCT ATACGCAGCA GCTGGCCTTC     1350
1351    AAACAACCTT CTTCAGCTTA TGCTGCCTTT AATAACCGTC CTCCTTCTAC     1400
1401    GTGGCTTACG GCCTACGTGG TCAAGGTATT TTTACTGGCA GCTAACCTCA     1450
1451    TTGCGATTGA TAGCCACGTG TTATGCGGCG CCGTTAAATG GTTGATTCTC     1500
1501    GAGAAGCAGA AGCCGGATGG AGTTTTTCAA GAAGACGGAC CGGTCATTCA     1550
1551    CCAAGAGATG ATTGGTGGTT TTCGCAACGC CAAGGAGGCA GATGTCTCAC     1600
1601    TGACGGCATT CGTGCTCATC GCGCTTCAAG AAGCACGTGA CATTTGCGAA     1650
1651    GGACAAGTAA ACAGCCTTCC CGGCTCCATT AATAAGGCTG GTGAGTACAT     1700
1701    TGAGGCGTCA TATATGAATC TTCAACGTCC TTATACGGTC GCTATCGCGG     1750
1751    GCTACGCCCT GGCCCTCATG AACAACTTG AGGAACCATA CCTAGGAAAA     1800
1801    TTCCTGAATA CAGCCAAGGA TCGTAATCGT TGGGAGGAGC CTGATCAGCA     1850

```

1851	GCTCTACAAC	GTGGAAGCCA	CTTCCTACGC	TCTTCTCGCA	CTGCTTCTCC	1900
1901	TGAAGGATTT	CGACTCCGTG	CCCCCTGTAG	TGCGCTGGCT	GAACGAACAA	1950
1951	CGTTACTATG	GGGGGGGGTA	TGGATCTACG	CAAGCAACAT	TCATGGTATT	2000
2001	TCAAGCCTTA	GCTCAGTATC	AGACTGATGT	ACCAGACCAC	AAGGATCTTA	2050
2051	ATATGGATGT	GTCCTTCCAC	CTCCCCTCAT	CAGGGTCCGG	AGGGGGTGGA	2100
2101	TCAGGGGGCG	GAGGATCTAC	GCCTGCCGGT	AGTGGTGAGC	AAAATATGAT	2150
2151	CGGGATGACC	CCTACTGTGA	TCGCCGTGCA	CTATTTAGAT	CAAACGGAGC	2200
2201	AATGGGAAAA	ATTTGGGATT	GAAAAACGTC	AGGAAGCGTT	AGAATTGATT	2250
2251	AAAAAGGGAT	ATACACAACA	ATTAGCGTTT	AAGCAACCAT	CAAGCGCGTA	2300
2301	TGCCGCGTTT	AACAACAGAC	CACCATCAAC	ATGGTTAACC	GCGTATGTCG	2350
2351	TGAAAGTGTT	TAGTTTGCGG	GCGAATTTAA	TTGCTATTGA	TAGTCACGTA	2400
2401	TTATGCGGAG	CGGTAAAGTG	GCTCATCTTA	GAAAAGCAAA	AACCAGACGG	2450
2451	CGTGTTCCAA	GAAGACGGAC	CAGTCATCCA	CCAGGAGATG	ATCGGGGGCT	2500
2501	TTAGAAATGC	GAAAGAAGCG	GACGTAAGCT	TAACCGCCTT	TGTATTGATT	2550
2551	GCCTTACAAG	AGGCGCGCGA	TATTTGCGAA	GGCCAAGTGA	ACTCTTTGCC	2600
2601	GGGATCGATT	AATAAAGCGG	GCGAATACAT	CGAGGCATCC	TATATGAATT	2650
2651	TACAACGCCC	TTATACCGTA	GCGATCGCCG	GATACGCGTT	AGCGTTAATG	2700
2701	AATAAGTTAG	AAGAGCCATA	TTTGGGGAAA	TTCTTAAATA	CGGCGAAGGA	2750
2751	CCGTAATAGG	TGGGAAGAAC	CAGATCAACA	ATTGTATAAT	GTCGAAGCGA	2800
2801	CCAGTTATGC	GTTGTTAGCG	TTATTACTTT	TAAAGGATTT	CGATAGCGTC	2850
2851	CCACCAGTGG	TCAGATGGTT	AAACGAACAG	CGCTATTATG	GCGGAGGTTA	2900
2901	CGGGAGTACA	CAAGCGACGT	TTATGGTCTT	TCAGGCGCTC	GCGCAGTACC	2950
2951	AGACGGACGT	GCCAGATCAC	AAAGACCTCA	ATATGGACGT	CAGTTTTCAC	3000
3001	TTGCCATCAT	CCGGGAGCGG	CGGAGGTGGG	AGCGGAGGGG	GCGGTACCTC	3050
3051	CGGATCTACC	CCCGCAGGCT	CTGGGGAACA	GAACATGATT	GGCATGACAC	3100
3101	CAACAGTCAT	TGCGGTACAC	TACCTGGACC	AGACCGAACA	GTGGGAGAAG	3150
3151	TTCGGCATAG	AGAAGAGGCA	AGAGGCCCTG	GAGCTCATCA	AGAAAGGGTA	3200
3201	CACCCAGCAG	CTGGCCTTCA	AACAGCCCAG	CTCTGCCATAT	GCTGCCTTCA	3250
3251	ACAACCGGCC	CCCCAGCACC	TGGCTGACAG	CCTACGTGGT	CAAGGTCCTC	3300
3301	TCTCTAGCTG	CCAACCTCAT	CGCCATCGAC	TCTCACGTCC	TGTGTGGGGC	3350
3351	TGTTAAATGG	TTGATTCTGG	AGAAACAGAA	GCCGGATGGT	GTCTTTCAGG	3400
3401	AGGATGGGCC	CGTGATTAC	CAAGAAATGA	TTGGTGGCTT	CCGGAACGCC	3450
3451	AAGGAGGCAG	ATGTGTCACT	CACAGCCTTC	GTCCTCATCG	CACTGCAGGA	3500
3501	AGCCAGGGAC	ATCTGTGAGG	GGCAGGTCAA	TAGCCTTCCT	GGGAGCATCA	3550
3551	ACAAGGCAGG	GGAGTATATT	GAAGCCAGTT	ACATGAACCT	GCAGAGACCA	3600
3601	TACACAGTGG	CCATTGCTGG	GTATGCCCTG	GCCCTGATGA	ACAAACTGGA	3650

3651 GGAACCTTAC CTCGGCAAGT TTCTGAACAC AGCCAAAGAT CGGAACCGCT 3700
3701 GGGAGGAGCC TGACCAGCAG CTCTACAACG TAGAGGCCAC ATCCTACGCC 3750
3751 CTCCTGGCCC TGCTGCTGCT GAAAGACTTT GACTCTGTGC CCCCTGTAGT 3800
3801 GCGCTGGCTC AATGAGCAAA GATACTACGG AGGCGGCTAT GGCTCCACCC 3850
3851 AGGCTACCTT CATGGTATTC CAAGCCTTGG CCCAATATCA AACAGATGTC 3900
3901 CCTGACCATA AGGACTTGAA CATGGATGTG TCCTTCCACC TCCCCAGCAG 3950
3951 TGGATCC... 3957

SEQ ID 33: Coding sequence from pVK96-02 (murine C3d₃-PyMSP1.19 in PVAX3 (antigen sequence underlined) from example 7).

```

706   . . . . .ATGGA TGCAATGAAG AGAGGGCTCT GCTGTGTGCT GCTGCTGTGT      750
751   GGAGCAGTCT TCGTTTCCGC TAGATCTACC CCAGCGGGCT CCGGAGAACA      800
801   AAACATGATT GGAATGACGC CTACAGTCAT TGCGGTCCAC TACCTGGACC      850
851   AGACCGAACA GTGGGAGAAA TTCGGAATCG AGAAACGCCA AGAAGCACTG      900
901   GAGCTGATTA AAAAGGGCTA TACGCAGCAG CTGGCCTTCA AACAACTTC      950
951   TTCAGCTTAT GCTGCCTTTA ATAACCGTCC TCCTTCTACG TGGCTTACGG     1000
1001  CCTACGTGGT CAAGGTATTT TCACTGGCAG CTAACCTCAT TGCGATTGAT     1050
1051  AGCCACGTGT TATGCGGCGC CGTTAAATGG TTGATTCTCG AGAAGCAGAA     1100
1101  GCCGGATGGA GTTTTTCAAG AAGACGGACC GGTCAATCAC CAAGAGATGA     1150
1151  TTGGTGGTTT TCGCAACGCC AAGGAGGCAG ATGTCTCACT GACGGCATTCT     1200
1201  GTGCTCATCG CGCTTCAAGA AGCACGTGAC ATTTGCGAAG GACAAGTAAA     1250
1251  CAGCCTTCCC GGCTCCATTA ATAAGGCTGG TGAGTACATT GAGGCGTCAT     1300
1301  ATATGAATCT TCAACGTCCT TATACGTCG CTATCGCGGG CTACGCCCTG     1350
1351  GCCCTCATGA ACAAACTTGA GGAACCATAC CTAGGAAAAT TCCTGAATAC     1400
1401  AGCCAAGGAT CGTAATCGTT GGGAGGAGCC TGATCAGCAG CTCTACAACG     1450
1451  TGGAAGCCAC TTCCTACGCT CTTCTCGCAC TGCTTCTCCT GAAGGATTTCT     1500
1501  GACTCCGTGC CCCCTGTAGT GCGCTGGCTG AACGAACAAC GTTACTATGG     1550
1551  GGGGGGGTAT GGATCTACGC AAGCAACATT CATGGTATTT CAAGCCTTAG     1600
1601  CTCAGTATCA GACTGATGTA CCAGACCACA AGGATCTTAA TATGGATGTG     1650
1651  TCCTTCCACC TCCCCTCATC AGGGTCCGGA GGGGGTGGAT CAGGGGGCGG     1700
1701  AGGATCTACG CCTGCCGGTA GTGGTGAGCA AAATATGATC GGGATGACCC     1750
1751  CTACTGTGAT CGCCGTGCAC TATTTAGATC AAACGGAGCA ATGGGAAAAA     1800
1801  TTTGGGATTG AAAAACGTCA GGAAGCGTTA GAATTGATTA AAAAGGGATA     1850
1851  TACACAACAA TTAGCGTTTA AGCAACCATC AAGCGCGTAT GCCGCGTTTA     1900
1901  ACAACAGACC ACCATCAACA TGGTAAACCG CGTATGTCGT GAAAGTGTTC     1950
1951  AGTTTGGCGG CGAATTTAAT TGCTATTGAT AGTCACGTAT TATGCGGAGC     2000
2001  GGTAAAGTGG CTCATCTTAG AAAAGCAAAA ACCAGACGGC GTGTTCCAAG     2050
2051  AAGACGGACC AGTCATCCAC CAGGAGATGA TCGGGGGCTT TAGAAATGCG     2100
2101  AAAGAAGCGG ACGTAAGCTT AACCGCCTTT GTATTGATTG CCTTACAAGA     2150
2151  GGCGCGCGAT ATTTGCGAAG GCCAAGTGAA CTCTTTGCCG GGATCGATTA     2200
2201  ATAAAGCGGG CGAATACATC GAGGCATCCT ATATGAATTT ACAACGCCCT     2250
2251  TATACCGTAG CGATCGCCGG ATACGCGTTA GCGTTAATGA ATAAGTTAGA     2300
2301  AGAGCCATAT TTGGGGAAAT TCTTAAATAC GGCGAAGGAC CGTAATAGGT     2350
2351  GGGAAGAACC AGATCAACAA TTGTATAATG TCGAAGCGAC CAGTTATGCG     2400
2401  TTGTTAGCGT TATTACTTTT AAAGGATTTT GATAGCGTCC CACCAGTGGT     2450
2451  CAGATGGTTA AACGAACAGC GCTATTATGG CGGAGGTTAC GGGAGTACAC     2500

```

2501 AAGCGACGTT TATGGTCTTT CAGGCGCTCG CGCAGTACCA GACGGACGTG 2550

2551	CCAGATCACA AAGACCTCAA TATGGACGTC AGTTTTTCACT TGCCATCATC	2600
2601	CGGGAGCGGC GGAGGTGGGA GCGGAGGGGG CGGTACCTCC GGATCTACCC	2650
2651	CCGCAGGCTC TGGGGAACAG AACATGATTG GCATGACACC AACAGTCATT	2700
2701	GCGGTACACT ACCTGGACCA GACCGAACAG TGGGAGAAGT TCGGCATAGA	2750
2751	GAAGAGGCAA GAGGCCCTGG AGCTCATCAA GAAAGGGTAC ACCCAGCAGC	2800
2801	TGGCCTTCAA ACAGCCCAGC TCTGCCTATG CTGCCTTCAA CAACCGGCCC	2850
2851	CCCAGCACCT GGCTGACAGC CTACGTGGTC AAGGTCTTCT CTCTAGCTGC	2900
2901	CAACCTCATC GCCATCGACT CTCACGTCTT GTGTGGGGCT GTTAAATGGT	2950
2951	TGATTCTGGA GAAACAGAAG CCGGATGGTG TCTTTCAGGA GGATGGGCCC	3000
3001	GTGATTCACC AAGAAATGAT TGGTGGCTTC CGGAACGCCA AGGAGGCAGA	3050
3051	TGTGTCACTC ACAGCCTTCG TCCTCATCGC ACTGCAGGAA GCCAGGGACA	3100
3101	TCTGTGAGGG GCAGGTCAAT AGCCTTCCTG GGAGCATCAA CAAGGCAGGG	3150
3151	GAGTATATTG AAGCCAGTTA CATGAACCTG CAGAGACCAT ACACAGTGGC	3200
3201	CATTGCTGGG TATGCCCTGG CCCTGATGAA CAAACTGGAG GAACCTTACC	3250
3251	TCGGCAAGTT TCTGAACACA GCCAAAGATC GGAACCGCTG GGAGGAGCCT	3300
3301	GACCAGCAGC TCTACAACGT AGAGGCCACA TCCTACGCCC TCCTGGCCCT	3350
3351	GCTGCTGCTG AAAGACTTTG ACTCTGTGCC CCCTGTAGTG CGCTGGCTCA	3400
3401	ATGAGCAAAG ATACTACGGA GGCGGCTATG GCTCCACCCA GGCTACCTTC	3450
3451	ATGGTATTCC AAGCCTTGGC CCAATATCAA ACAGATGTCC CTGACCATAA	3500
3501	GGACTTGAAC ATGGATGTGT CCTTCCACCT CCCCAGCAGT GGATCTCACA	3550
3551	<u>TTGCCTCTAT TGCTTTGAAC AACTTGAACA AGTCTGGTTT GGTAGGAGAA</u>	3600
3601	<u>GGTGAGTCTA AGAAGATTTT GGCTAAGATG CTGAACATGG ACGGTATGGA</u>	3650
3651	<u>CTTGTTGGGT GTTGACCCTA AGCATGTTTG TGTTGACACT AGAGACATTC</u>	3700
3701	<u>CTAAGAACGC TGGATGTTTC AGAGACGACA ACGGTACTGA AGAGTGGAGA</u>	3750
3751	<u>TGTTTGTTGG GTTACAAGAA GGGTGAGGGT AACACCTGCG TTGAGAACAA</u>	3800
3801	<u>CAACCCTACT TGCACATCA ACAACGGTGG ATGTGACCCA ACCGCCTCTT</u>	3850
3851	<u>GTCAAAACGC TGAATCTACC GAAAACCTCCA AGAAGATTAT TTGCACCTGT</u>	3900
3901	<u>AAGGAACCAA CCCCTAACGC CTAACGAG GGTGTTTTCT GTTCTTCTTC</u>	3950
3951	<u>CGGATCC...</u>	3957

SEQ ID 34:

DNA sequence of oligonucleotide used for Reverse Transcriptase priming in Example 9

1 GGGGGCAGTT GGAGGGACAC ATCAAG 26

SEQ ID 35:

DNA sequence of PCR oligonucleotide primer used in Example 9

1 GGACCCCTC GGGCTGCGG GAAC 24

SEQ ID 36:

DNA sequence of PCR oligonucleotide primer used in Example 9

1 GGCATATGAC CCCGTCGGGC TCCGGGGAA 29

SEQ ID 37:

DNA sequence of PCR oligonucleotide primer used in Example 9

1 GGAATTCAGC AGGATCCACT GCTGGGCAGT TGGAGGGACA CATCAAG 47

SEQ ID 38:

DNA sequence of wild-type human C3d (Example 9)

1	ACCCCTCGG GCTCCGGGGA ACAGAACATG ATCGGCATGA CGCCACGGT	50
51	CATCGCTGTG CATTACCTGG ATGAAACGGA GCAGTGGGAG AAGTTCGGCC	100
101	TAGAGAAGCG GCAGGGGGCC TTGGAGCTCA TCAAGAAGGG GTACACCCAG	150
151	CAGCTGGCCT TCAGACAACC CAGCTCTGCC TTTGCGGCCT TCGTGAAACG	200
201	GGCACCAGC ACCTGGCTGA CCGCCTACGT GGTCAAGGTC TTCTCTCTGG	250
251	CTGTCAACCT CATCGCCATC GACTCCCAAG TCCTCTGCGG GGCTGTAAAA	300
301	TGGCTGATCC TGGAGAAGCA GAAGCCCGAC GGGGTCTTCC AGGAGGATAC	350
351	GCCCGTGATA CACCAAGAAA TGATTGGTGG ATTACGGAAC AACAAACGAGA	400
401	AAGACATGGC CCTCACGGCC TTTGTTCTCA TCTCGCTGCA GGAGGCTAAA	450
451	GATATTTGCG AGGAGCAGGT CAACAGCCTG CCAGGCAGCA TCACTAAAGC	500
501	AGGAGACTTC CTTGAAGCCA ACTACATGAA CCTACAGAGA TCCTACACTG	550
551	TGGCCATTGC TGGCTATGCT CTGGCCCAGA TGGGCAGGCT GAAGGGGCCT	600
601	CTTCTTAACA AATTTCTGAC CACAGCCAAA GATAAGAACC GCTGGGAGGA	650
651	CCCTGGTAAG CAGCTCTACA ACGTGGAGGC CACATCCTAT GCCCTCTTGG	700
701	CCCTACTGCA ACTAAAAGAC TTTGACTTTG TGCCTCCCGT CGTGCCTTGG	750
751	CTCAATGAAC AGAGATACTA CGGTGGTGGC TATGGCTCTA CCCAGGCCAC	800
801	CTTCATGGTG TTCCAAGCCT TGGCTCAATA CAAAAGGAC GCCCCTGACC	850
851	ACCAGGAACT GAACCTTGAT GTGTCCCTCC AACTGCCCTG A	891

SEQ ID 39:

DNA sequence of first human C3d variant (Example 10)

1	CCAGATCTAC	GCCAAGCGGA	TCAGGCGAGC	AGAATATGAT	CGGGATGACA	50
51	CCAACCGTAA	TTGCGGTCCA	TTATCTCGAC	GAAACCGAAC	AGTGGGAAAA	100
101	ATTTGGGCTC	GAAAAGCGTC	AAGGCGCTCT	CGAGTTGATC	AAGAAAGGCT	150
151	ACACGCAACA	GTTAGCGTTC	CGTCAACCAT	CATCAGCGTT	CGCCGCTTTC	200
201	GTAAAGCGTG	CGCCATCAAC	GTGGCTCACA	GCGTATGTAG	TGAAGGTATT	250
251	TAGCCTCGCC	GTAAATTTAA	TCGCGATTGA	CAGTCAAGTG	TTATGCGGCG	300
301	CGGTCAAGTG	GCTCATTCTT	GAAAAGCAAA	AGCCAGATGG	CGTATTCCAA	350
351	GAGGACGCCC	CAGTCATCCA	CCAAGAGATG	ATTGGCGGCC	TCCGCAATAA	400
401	CAATGAGAAG	GACATGGCGT	TAACCGCGTT	TGTCTTAATC	AGTTTACAGG	450
451	AAGCCAAAGA	CATTTGTGAG	GAACAGGTAA	ATAGTTTACC	TGGGAGTATT	500
501	ACGAAAGCGG	GCGATTTCTT	AGAAGCAAAT	TACATGAATC	TCCAACGCTC	550
551	ATACACGGTA	GCGATCGCGG	GATATGCCTT	AGCGCAGATG	GGGAGATTAA	600
601	AAGGCCCAT	ACTGAACAAG	TTTTTAACAA	CCGCAAAAGA	CAAGAATAGG	650
651	TGGGAGGACC	CAGGCAAGCA	ACTTTATAAC	GTCGAAGCAA	CGTCATACGC	700
701	ATTATTAGCA	CTCTTACAAC	TCAAGGACTT	CGACTTCGTA	CCACCTGTGG	750
751	TACGGTGGCT	TAACGAACAA	AGGTATTACG	GGGGCGGATA	CGGCAGCACG	800
801	CAAGCGACTT	TCATGGTCTT	TCAAGCACTC	GCACAGTACC	AGAAGGATGC	850
851	ACCTGATCAC	CAAGAATTAA	ACTTAGATGT	CAGTCTGCAG	TTACCAAGTT	900
901	CAGGGTCAGG	TGGAGGTGGA	AGTGGTGGAG	GTGGAAGCGG	ATCCTAA	947

SEQ ID 40:

DNA sequence of second human C3d variant (Example 10)

1	CCAGATCTAC	TCCTTCAGGG	AGTGGAGAAC	AAAACATGAT	TGGTATGACC	50
51	CCTACAGTGA	TCGCCGTACA	CTACTTAGAT	GAGACAGAGC	AATGGGAGAA	100
101	ATTTCGGTTT	GAGAAAAGAC	AGGGAGCGTT	AGAACTTATT	AAAAAGGGAT	150
151	ATACACAGCA	ACTCGCTTTT	AGGCAGCCTA	GTAGCGCATT	TGCTGCGTTT	200
201	GTCAAAAGAG	CCCCTAGTAC	ATGGTTAACG	GCTTACGTCG	TAAAAGTGTT	250
251	CTCATTAGCG	GTGAACCTGA	TTGCAATCGA	TTCGCAGGTA	CTGTGTGGAG	300
301	CCGTGAAATG	GTTAATCTTA	GAGAAACAGA	AACCTGACGG	AGTGTTCAG	350
351	GAAGATGCAC	CTGTAATGCA	GCAGGAAATG	ATCGGGGGCT	TGACAAACAA	400
401	TAACGAAAAA	GATATGGCTC	TGACAGCTTT	CGTGCTGATT	TCCCTCCAAG	450
451	AGGCGAAGGA	TATCTGCGAA	GAGCAAGTGA	ACTCACTCCC	AGGATCAATC	500
501	ACCAAGGCCG	GGGACTTTCT	GGAGGCGAAC	TATATGAACT	TGCAGAGGAG	550
551	CTATACCGTC	GCAATTGCCG	GTTACGCGCT	CGCACAAATG	GGACGTCTCA	600
601	AAGGACCTCT	GTTAAATAAA	TTCCTCACGA	CGGCGAAGGA	TAAAACCGA	650
651	TGGGAAGACC	CTGGGAAACA	GTTGTACAAT	GTAGAGGCGA	CCAGTTATGC	700
701	GCTGCTCGCG	TTGCTCCAGT	TGAAAGATTT	TGATTTTGTC	CCTCCAGTAG	750
751	TCAGATGGTT	GAATGAGCAG	CGTTACTATG	GAGGGGGGTA	TGGATCAACA	800
801	CAGGCAACGT	TTATGGTATT	CCAGGCGTTA	GCGCAATATC	AAAAAGACGC	850
851	GCCAGACCAC	CAGGAGCTTA	ATCTCGACGT	ATCATTACAA	CTCCCTTCAA	900
901	GCGGCAGCGG	CGGGGGCGGG	TCAGGAGGCG	GGGGTTCTGG	ATCCTAA	947

SEQ ID 41:

DNA sequence of concatamer of two variant human C3d sequences and one wild-type sequence for baculovirus expression vector pBP80-01 (Example 11)

1	ATGGCCCTCT	GGATGCGCCT	CCTGCCCCCTG	CTGGCCCTGC	TGGCCCTCTG	50
51	GGCGCCCGCG	CCCACCCGAG	CCGGATCCAG	ATCTACGCCA	AGCGGATCAG	100
101	GCGAGCAGAA	TATGATCGGG	ATGACACCAA	CCGTAATTGC	GGTCCATTAT	150
151	CTCGACGAAA	CCGAACAGTG	GGAAAAATT	GGGCTCGAAA	AGCGTCAAGG	200
201	CGCTCTCGAG	TTGATCAAGA	AAGGCTACAC	GCAACAGTTA	GCGTTCGTC	250
251	AACCATCATC	AGCGTTCGCC	GCTTTCGTAA	AGCGTGCGCC	ATCAACGTGG	300
301	CTCACAGCGT	ATGTAGTGAA	GGTATTTAGC	CTCGCCGTAA	ATTTAATCGC	350
351	GATTGACAGT	CAAGTGTTAT	GCGGCGCGGT	CAAGTGGCTC	ATTCTTGAAA	400
401	AGCAAAAGCC	AGATGGCGTA	TTCCAAGAGG	ACGCCCCAGT	CATCCACCAA	450
451	GAGATGATTG	GCGGCCTCCG	CAATAACAAT	GAGAAGGACA	TGGCGTTAAC	500
501	CGCGTTTGTC	TTAATCAGTT	TACAGGAAGC	CAAAGACATT	TGTGAGGAAC	550
551	AGGTAAATAG	TTTACCTGGG	AGTATTACGA	AAGCGGGCGA	TTTCTTAGAA	600
601	GCAAATTACA	TGAATCTCCA	ACGCTCATAC	ACGGTAGCGA	TCGCGGGATA	650
651	TGCCTTAGCG	CAGATGGGGA	GATTAAAAGG	CCCATTACTG	AACAAGTTT	700
701	TAACAACCGC	AAAAGACAAG	AATAGGTGGG	AGGACCCAGG	CAAGCAACTT	750
751	TATAACGTCG	AAGCAACGTC	ATACGCATTA	TTAGCACTCT	TACAACTCAA	800
801	GGACTTCGAC	TTCGTACCAC	CTGTGGTACG	GTGGCTTAAC	GAACAAAGGT	850
851	ATTACGGGGG	CGGATACGGC	AGCACGCAAG	CGACTTTCAT	GGTCTTTCAA	900
901	GCACTCGCAC	AGTACCAGAA	GGATGCACCT	GATCACCAAG	AATTAAACTT	950
951	AGATGTCAGT	CTGCAGTTAC	CAAGTTCAGG	GTCAGGTGGA	GGTGGAAGTG	1000
1001	GTGGAGGTGG	AAGCGGATCT	ACTCCTTCAG	GGAGTGGAGA	ACAAAACATG	1050
1051	ATTGGTATGA	CCCCTACAGT	GATCGCCGTA	CACTACTTAG	ATGAGACAGA	1100
1101	GCAATGGGAG	AAATTCGGTT	TGGAGAAAAG	ACAGGGAGCG	TTAGAACTTA	1150
1151	TTAAAAAGGG	ATATACACAG	CAACTCGCTT	TTAGGCAGCC	TAGTAGCGCA	1200
1201	TTTGCTGCGT	TTGTCAAAAG	AGCCCCTAGT	ACATGGTTAA	CGGCTTACGT	1250
1251	CGTAAAAGTG	TTCTCATTAG	CGGTGAACCT	GATTGCAATC	GATTCGCAGG	1300
1301	TACTGTGTGG	AGCCGTGAAA	TGGTTAATCT	TAGAGAAACA	GAAACCTGAC	1350
1351	GGAGTGT TTC	AGGAAGATGC	ACCTGTAATT	CACCAGGAAA	TGATCGGGGG	1400
1401	CTTGAGAAAC	AATAACGAAA	AAGATATGGC	TCTGACAGCT	TTCTGTCTGA	1450
1451	TTTCCCTCCA	AGAGGCGAAG	GATATCTGCG	AAGAGCAAGT	GAACCTCACTC	1500
1501	CCAGGATCAA	TCACCAAGGC	CGGGGACTTT	CTGGAGGCGA	ACTATATGAA	1550
1551	CTTGCAGAGG	AGCTATACCG	TCGCAATTGC	CGGTTACGCG	CTCGCACAAA	1600
1601	TGGGACGTCT	CAAAGGACCT	CTGTTAAATA	AATTCCTCAC	GACGGCGAAG	1650
1651	GATAAAAACC	GATGGGAAGA	CCCTGGGAAA	CAGTTGTACA	ATGTAGAGGC	1700

1701	GACCAGTTAT	GCGCTGCTCG	CGTTGCTCCA	GTTGAAAGAT	TTTGATTTTG	1750
1751	TCCCTCCAGT	AGTCAGATGG	TTGAATGAGC	AGCGTTACTA	TGGAGGGGGG	1800
1801	TATGGATCAA	CACAGGCAAC	GTTTATGGTA	TTCCAGGCGT	TAGCGCAATA	1850
1851	TCAAAAAGAC	GCGCCAGACC	ACCAGGAGCT	TAATCTCGAC	GTATCATTAC	1900
1901	AACTCCCTTC	AAGCGGCAGC	GGCGGGGGCG	GGTCAGGAGG	CGGGGGTTCT	1950
1951	GGATCTACCC	CCTCGGGCTC	CGGGGAACAG	AACATGATCG	GCATGACGCC	2000
2001	CACGGTCATC	GCTGTGCATT	ACCTGGATGA	AACGGAGCAG	TGGGAGAAAGT	2050
2051	TCGGCCTAGA	GAAGCGGCAG	GGGGCCTTGG	AGCTCATCAA	GAAGGGGTAC	2100
2101	ACCCAGCAGC	TGGCCTTCAG	ACAACCCAGC	TCTGCCTTTG	CGGCCTTCGT	2150
2151	GAAACGGGCA	CCCAGCACCT	GGCTGACCGC	CTACGTGGTC	AAGGTCTTCT	2200
2201	CTCTGGCTGT	CAACCTCATC	GCCATCGACT	CCCAAGTCCT	CTGCGGGGCT	2250
2251	GTTAAATGGC	TGATCCTGGA	GAAGCAGAAG	CCCGACGGGG	TCTTCCAGGA	2300
2301	GGATGCGCCC	GTGATACACC	AAGAAATGAT	TGGTGGATTA	CGGAACAACA	2350
2351	ACGAGAAAGA	CATGGCCCTC	ACGGCCTTTG	TTCTCATCTC	GCTGCAGGAG	2400
2401	GCTAAAGATA	TTTGCGAGGA	GCAGGTCAAC	AGCCTGCCAG	GCAGCATCAC	2450
2451	TAAAGCAGGA	GACTTCCTTG	AAGCCAACTA	CATGAACCTA	CAGAGATCCT	2500
2501	ACACTGTGGC	CATTGCTGGC	TATGCTCTGG	CCCAGATGGG	CAGGCTGAAG	2550
2551	GGGCCTCTTC	TTAACAAATT	TCTGACCACA	GCCAAAGATA	AGAACCGCTG	2600
2601	GGAGGACCCT	GGTAAGCAGC	TCTACAACGT	GGAGGCCACA	TCCTATGCCC	2650
2651	TCTTGGCCCT	ACTGCAGCTA	AAAGACTTTG	ACTTTGTGCC	TCCCGTCGTG	2700
2701	CGTTGGCTCA	ATGAACAGAG	ATACTACGGT	GGTGGCTATG	GCTCTACCCA	2750
2751	GGCCACCTTC	ATGGTGTTC	AAGCCTTGGC	TCAATACCAA	AAGGACGCCC	2800
2801	CTGACCACCA	GGAAGTGAAC	CTTGATGTGT	CCCTCCAAC	GCCCAGCAGT	2850
2851	GGATCCTGCT	GACTCGAGGC	CTGCAGGGCG	GCCGCTTAAT	TAATTGA	2897

SEQ ID 42: Coding sequence from pVK80-01 (human C3d₃ in PVAX3 (signal peptide in italics) from example 13).

706ATGGA	TGCAATGAAG	AGAGGGCTCT	GCTGTGTGCT	GCTGCTGTGT	750
751	GGAGCAGTCT	TCGTTTCCGC	TAGATCTACG	CCAAGCGGAT	CAGGCGAGCA	800
801	GAATATGATC	GGGATGACAC	CAACCGTAAT	TGCGGTCCAT	TATCTCGACG	850
851	AAACCGAACA	GTGGGAAAAA	TTTGGGCTCG	AAAAGCGTCA	AGGCGCTCTC	900
901	GAGTTGATCA	AGAAAGGCTA	CACGCAACAG	TTAGCGTTCC	GTCAACCATC	950
951	ATCAGCGTTC	GCCGCTTTTCG	TAAAGCGTGC	GCCATCAACG	TGGCTCACAG	1000
1001	CGTATGTAGT	GAAGGTATTT	AGCCTCGCCG	TAAATTTAAT	CGCGATTGAC	1050
1051	AGTCAAGTGT	TATGCGGCGC	GGTCAAGTGG	CTCATTCCTG	AAAAGCAAAA	1100
1101	GCCAGATGGC	GTATTCCAAG	AGGACGCCCC	AGTCATCCAC	CAAGAGATGA	1150
1151	TTGGCGGCCCT	CCGCAATAAC	AATGAGAAGG	ACATGGCGTT	AACCGCGTTT	1200
1201	GTCTTAATCA	GTTTACAGGA	AGCCAAAGAC	ATTTGTGAGG	AACAGGTAAA	1250
1251	TAGTTTACCT	GGGAGTATTA	CGAAAGCGGG	CGATTTCTTA	GAAGCAAATT	1300
1301	ACATGAATCT	CCAACGCTCA	TACACGGTAG	CGATCGCGGG	ATATGCCTTA	1350
1351	GCGCAGATGG	GGAGATTAAA	AGGCCCATTA	CTGAACAAGT	TTTTAACAAC	1400
1401	CGCAAAGAC	AAGAATAGGT	GGGAGGACCC	AGGCAAGCAA	CTTTATAACG	1450
1451	TCGAAGCAAC	GTCATACGCA	TTATTAGCAC	TCTTACAAC	CAAGGACTTC	1500
1501	GACTTCGTAC	CACCTGTGGT	ACGGTGGCTT	AACGAACAAA	GGTATTACGG	1550
1551	GGGCGGATAC	GGCAGCACGC	AAGCGACTTT	CATGGTCTTT	CAAGCACTCG	1600
1601	CACAGTACCA	GAAGGATGCA	CCTGATCACC	AAGAATTAAA	CTTAGATGTC	1650
1651	AGTCTGCAGT	TACCAAGTTC	AGGGTCAGGT	GGAGGTGGAA	GTGGTGGAGG	1700
1701	TGGAAGCGGA	TCTACTCCTT	CAGGGAGTGG	AGAACAAAAC	ATGATTGGTA	1750
1751	TGACCCCTAC	AGTGATCGCC	GTACACTACT	TAGATGAGAC	AGAGCAATGG	1800
1801	GAGAAATTCG	GTTTGGAGAA	AAGACAGGGA	GCGTTAGAAC	TTATTAAAAA	1850
1851	GGGATATACA	CAGCAACTCG	CTTTTAGGCA	GCCTAGTAGC	GCATTTGCTG	1900
1901	CGTTTGTCAA	AAGAGCCCCT	AGTACATGGT	TAACGGCTTA	CGTCGTAAAA	1950
1951	GTGTTCTCAT	TAGCGGTGAA	CCTGATTGCA	ATCGATTGCG	AGGTACTGTG	2000
2001	TGGAGCCGTG	AAATGGTTAA	TCTTAGAGAA	ACAGAAACCT	GACGGAGTGT	2050
2051	TTCAGGAAGA	TGCACCTGTA	ATTACCAGG	AAATGATCGG	GGGCTTGAGA	2100
2101	AACAATAACG	AAAAAGATAT	GGCTCTGACA	GCTTTCGTGC	TGATTTCCCT	2150
2151	CCAAGAGGCG	AAGGATATCT	GCGAAGAGCA	AGTGAAC	CTCCAGGAT	2200
2201	CAATCACCAA	GGCCGGGGAC	TTTCTGGAGG	CGAACTATAT	GAACCTGCAG	2250
2251	AGGAGCTATA	CCGTCGCAAT	TGCCGGTTAC	GCGCTCGCAC	AAATGGGACG	2300
2301	TCTCAAAGGA	CCTCTGTTAA	ATAAATTCCT	CACGACGGCG	AAGGATAAAA	2350
2351	ACCGATGGGA	AGACCCTGGG	AAACAGTTGT	ACAATGTAGA	GGCGACCAGT	2400
2401	TATGCGCTGC	TCGCGTTGCT	CCAGTTGAAA	GATTTTGATT	TTGTCCCTCC	2450
2451	AGTAGTCAGA	TGGTTGAATG	AGCAGCGTTA	CTATGGAGGG	GGGTATGGAT	2500

2501 CAACACAGGC AACGTTTATG GTATTCCAGG CGTTAGCGCA ATATCAAAAA 2550

2551	GACGCGCCAG	ACCACCAGGA	GCTTAATCTC	GACGTATCAT	TACAACTCCC	2600
2601	TTCAAGCGGC	AGCGGCGGGG	GCGGGTCAGG	AGGCGGGGGT	TCTGGATCTA	2650
2651	CCCCCTCGGG	CTCCGGGGAA	CAGAACATGA	TCGGCATGAC	GCCCACGGTC	2700
2701	ATCGCTGTGC	ATTACCTGGA	TGAAACGGAG	CAGTGGGAGA	AGTTCGGCCT	2750
2751	AGAGAAGCGG	CAGGGGGCCT	TGGAGCTCAT	CAAGAAGGGG	TACACCCAGC	2800
2801	AGCTGGCCTT	CAGACAACCC	AGCTCTGCCT	TTGCGGCCTT	CGTGAAACGG	2850
2851	GCACCCAGCA	CCTGGCTGAC	CGCCTACGTG	GTCAAGGTCT	TCTCTCTGGC	2900
2901	TGTCAACCTC	ATCGCCATCG	ACTCCCAAGT	CCTCTGCGGG	GCTGTTTAAAT	2950
2951	GGCTGATCCT	GGAGAAGCAG	AAGCCCGACG	GGGTCTTCCA	GGAGGATGCG	3000
3001	CCCGTGATAC	ACCAAGAAAT	GATTGGTGGG	TTACGGAACA	ACAACGAGAA	3050
3051	AGACATGGCC	CTCACGGCCT	TTGTTCTCAT	CTCGCTGCAG	GAGGCTAAAG	3100
3101	ATATTTGCGA	GGAGCAGGTC	AACAGCCTGC	CAGGCAGCAT	CACTAAAGCA	3150
3151	GGAGACTTCC	TTGAAGCCAA	CTACATGAAC	CTACAGAGAT	CCTACACTGT	3200
3201	GGCCATTGCT	GGCTATGCTC	TGGCCCAGAT	GGGCAGGCTG	AAGGGGCCTC	3250
3251	TTCTTAACAA	ATTTCTGACC	ACAGCCAAAG	ATAAGAACCG	CTGGGAGGAC	3300
3301	CCTGGTAAGC	AGCTCTACAA	CGTGGAGGCC	ACATCCTATG	CCCTCTTGGC	3350
3351	CCTACTGCAG	CTAAAAGACT	TTGACTTTGT	GCCTCCCGTC	GTGCGTTGGC	3400
3401	TCAATGAACA	GAGATACTAC	GGTGGTGGCT	ATGGCTCTAC	CCAGGCCACC	3450
3451	TTCATGGTGT	TCCAAGCCTT	GGCTCAATAC	CAAAAGGACG	CCCCTGACCA	3500
3501	CCAGGAAGTG	AACCTTGATG	TGTCCCTCCA	ACTGCCCAGC	AGTGGATCC.	3549

SEQ ID 43: DNA sequence of PfMSP1.19 insert from example 13).

1	AGATCTAACA	TTGCCCAACA	CCAATGCGTT	AAGAAGCAAT	GTCCACAAAA	50
51	CTCCGGATGT	TTCAGACATC	TGGACGAGAG	AGAAGAATGT	AAGTGTCTGT	100
101	TGAACTACAA	GCAGGAAGGT	GATAAGTGTG	TTGAGAACCC	AAACCCTACC	150
151	TGTAACGAGA	ACAACGGTGG	ATGCGACGCT	GACGCTAAGT	GCACCGAAGA	200
201	AGACTCTGGT	TCTAACGGAA	AGAAGATTAC	TTGCGAATGT	ACTAAGCCAG	250
251	ACTCTTACCC	TTTGTTCGAT	GGAATCTTCT	GTTCTTCCTC	TAACTCTTCC	300
301	TCTGGATCC					309

SEQ ID 44: Coding sequence from pVK104-01 (PfMSP1.19-human C3d₃ in PVAX3 (signal peptide in *italics*, antigen sequence underlined) from example 13).

706ATGGA	TGCAATGAAG	AGAGGGCTCT	GCTGTGTGCT	GCTGCTGTGT	750
751	GGAGCAGTCT	TCGTTTCCGC	TAGATCTAAC	ATTGCCCAAC	ACCAATGCGT	800
801	TAAGAAGCAA	TGTCCACAAA	ACTCCGGATG	TTTCAGACAT	CTGGACGAGA	850
851	GAGAAGAATG	TAAGTGTCTG	TTGAACTACA	AGCAGGAAGG	TGATAAGTGT	900
901	GTTGAGAACC	CAAACCTTAC	CTGTAACGAG	AACAACGGTG	GATGCGACGC	950
951	TGACGCTAAG	TGCACCGAAG	AAGACTCTGG	TTCTAACGGA	AAGAAGATTA	1000
1001	CTTGCGAATG	TACTAAGCCA	GACTCTTACC	CTTTGTTCTGA	TGGAATCTTC	1050
1051	TGTTCTTCCT	CTAACTCTTC	CTCTGGATCT	ACGCCAAGCG	GATCAGGCGA	1100
1101	GCAGAATATG	ATCGGGATGA	CACCAACCGT	AATTGCGGTC	CATTATCTCG	1150
1151	ACGAAACCGA	ACAGTGGGAA	AAATTTGGGC	TCGAAAAGCG	TCAAGGCGCT	1200
1201	CTCGAGTTGA	TCAAGAAAGG	CTACACGCAA	CAGTTAGCGT	TCCGTCAACC	1250
1251	ATCATCAGCG	TTCGCCGCTT	TCGTAAAGCG	TGCGCCATCA	ACGTGGCTCA	1300
1301	CAGCGTATGT	AGTGAAGGTA	TTTAGCCTCG	CCGTAAATTT	AATCGCGATT	1350
1351	GACAGTCAAG	TGTTATGCGG	CGCGGTCAAG	TGGCTCATTC	TTGAAAAGCA	1400
1401	AAAGCCAGAT	GGCGTATTCC	AAGAGGACGC	CCCAGTCATC	CACCAAGAGA	1450
1451	TGATTGGCGG	CCTCCGCAAT	AACAATGAGA	AGGACATGGC	GTTAACC CGC	1500
1501	TTTGTCTTAA	TCAGTTTACA	GGAAGCCAAA	GACATTTGTG	AGGAACAGGT	1550
1551	AAATAGTTTA	CCTGGGAGTA	TTACGAAAGC	GGGCGATTTT	TTAGAAGCAA	1600
1601	ATTACATGAA	TCTCCAACGC	TCATACACGG	TAGCGATCGC	GGGATATGCC	1650
1651	TTAGCGCAGA	TGGGGAGATT	AAAAGGCCCA	TTACTGAACA	AGTTTTTAAC	1700
1701	AACCGCAAAA	GACAAGAATA	GGTGGGAGGA	CCCAGGCAAG	CAACTTTATA	1750
1751	ACGTCGAAGC	AACGTCATAC	GCATTATTAG	CACTCTTACA	ACTCAAGGAC	1800
1801	TTGACTTTCG	TACCACCTGT	GGTACGGTGG	CTTAACGAAC	AAAGGTATTA	1850
1851	CGGGGGCGGA	TACGGCAGCA	CGCAAGCGAC	TTTCATGGTC	TTTCAAGCAC	1900
1901	TCGCACAGTA	CCAGAAGGAT	GCACCTGATC	ACCAAGAATT	AAACTTAGAT	1950
1951	GTCAGTCTGC	AGTTACCAAG	TTCAGGGTCA	GGTGGAGGTG	GAAGTGGTGG	2000
2001	AGGTGGAAGC	GGATCTACTC	CTTCAGGGAG	TGGAGAACAA	AACATGATTG	2050
2051	GTATGACCCC	TACAGTGATC	GCCGTACACT	ACTTAGATGA	GACAGAGCAA	2100
2101	TGGGAGAAAT	TCGGTTTGGA	GAAAAGACAG	GGAGCGTTAG	AACTTATTAA	2150
2151	AAAGGGATAT	ACACAGCAAC	TCGCTTTTAG	GCAGCCTAGT	AGCGCATTTG	2200
2201	CTGCGTTTGT	CAAAAGAGCC	CCTAGTACAT	GGTTAACGGC	TTACGTCGTA	2250
2251	AAAGTGTTCT	CATTAGCGGT	GAACCTGATT	GCAATCGATT	CGCAGGTA	2300
2301	GTGTGGAGCC	GTGAAATGGT	TAATCTTAGA	GAAACAGAAA	CCTGACGGAG	2350
2351	TGTTTCAGGA	AGATGCACCT	GTAATTCACC	AGGAAATGAT	CGGGGGCTTG	2400
2401	AGAAACAATA	ACGAAAAAGA	TATGGCTCTG	ACAGCTTTCG	TGCTGATTTT	2450

2451 CCTCCAAGAG GCGAAGGATA TCTGCGAAGA GCAAGTGAAC TCACTCCCAG 2500
2501 GATCAATCAC CAAGGCCGGG GACTTTCTGG AGGCGAACTA TATGAACTTG 2550

2551	CAGAGGAGCT	ATACCGTCGC	AATTGCCGGT	TACGCGCTCG	CACAAATGGG	2600
2601	ACGTCTCAAA	GGACCTCTGT	TAAATAAAAT	CCTCACGACG	GCGAAGGATA	2650
2651	AAAACCGATG	GGAAGACCCT	GGGAAACAGT	TGTACAATGT	AGAGGCGACC	2700
2701	AGTTATGCGC	TGCTCGCGTT	GCTCCAGTTG	AAAGATTTTG	ATTTTGTCCC	2750
2751	TCCAGTAGTC	AGATGGTTGA	ATGAGCAGCG	TTACTATGGA	GGGGGGTATG	2800
2801	GATCAACACA	GGCAACGTTT	ATGGTATTCC	AGGCGTTAGC	GCAATATCAA	2850
2851	AAAGACGCGC	CAGACCACCA	GGAGCTTAAT	CTCGACGTAT	CATTACAAC	2900
2901	CCCTTCAAGC	GGCAGCGGCG	GGGGCGGGTC	AGGAGGCGGG	GGTTCTGGAT	2950
2951	CTACCCCTC	GGGCTCCGGG	GAACAGAACA	TGATCGGCAT	GACGCCCACG	3000
3001	GTCATCGCTG	TGCATTACCT	GGATGAAACG	GAGCAGTGGG	AGAAGTTCGG	3050
3051	CCTAGAGGAG	CGGCAGGGGG	CCTTGGAGCT	CATCAAGAAG	GGGTACACCC	3100
3101	AGCAGCTGGC	CTTCAGACAA	CCCAGCTCTG	CCTTTGCGGC	CTTCGTGAAA	3150
3151	CGGGCACCCA	GCACCTGGCT	GACCGCCTAC	GTGGTCAAGG	TCTTCTCTCT	3200
3201	GGCTGTCAAC	CTCATCGCCA	TCGACTCCCA	AGTCCTCTGC	GGGGCTGTTA	3250
3251	AATGGCTGAT	CCTGGAGAAG	CAGAAGCCCG	ACGGGGTCTT	CCAGGAGGAT	3300
3301	GCGCCCGTGA	TACACCAAGA	AATGATTGGT	GGATTACGGA	ACAACAACGA	3350
3351	GAAAGACATG	GCCCTCACGG	CCTTTGTTCT	CATCTCGCTG	CAGGAGGCTA	3400
3401	AAGATATTTG	CGAGGAGCAG	GTCAACAGCC	TGCCAGGCAG	CATCACTAAA	3450
3451	GCAGGAGACT	TCCTTGAAGC	CAACTACATG	AACCTACAGA	GATCCTACAC	3500
3501	TGTGGCCATT	GCTGGCTATG	CTCTGGCCCA	GATGGGCAGG	CTGAAGGGGC	3550
3551	CTCTTCTTAA	CAAATTTCTG	ACCACAGCCA	AAGATAAGAA	CCGCTGGGAG	3600
3601	GACCCCTGGTA	AGCAGCTCTA	CAACGTGGAG	GCCACATCCT	ATGCCCTCTT	3650
3651	GGCCCTACTG	CAGCTAAAAG	ACTTTGACTT	TGTGCCTCCC	GTCGTGCGTT	3700
3701	GGCTCAATGA	ACAGAGATAC	TACGGTGGTG	GCTATGGCTC	TACCCAGGCC	3750
3751	ACCTTCATGG	TGTTCCAAGC	CTTGGCTCAA	TACCAAAAAGG	ACGCCCCTGA	3800
3801	CCACCAGGAA	CTGAACCTTG	ATGTGTCCCT	CCAAC TGCCC	AGCAGTGGAT	3850
3851	CC...					3852

SEQ ID 45: Coding sequence from pVK104-02 (human C3d₃-PfMSP1.19 in PVAX3 (signal peptide in *italics*, antigen sequence underlined) from example 13).

```

706  ....ATGGA TGCAATGAAG AGAGGGCTCT GCTGTGTGCT GCTGCTGTGT 750
751  GGAGCAGTCT TCGTTTCCGC TAGATCTACG CCAAGCGGAT CAGGCGAGCA 800
801  GAATATGATC GGGATGACAC CAACCGTAAT TCGGGTCCAT TATCTCGACG 850
851  AAACCGAACA GTGGGAAAAA TTTGGGCTCG AAAAGCGTCA AGGCGCTCTC 900
901  GAGTTGATCA AGAAAGGCTA CACGCAACAG TTAGCGTTCC GTCAACCATC 950
951  ATCAGCGTTC GCCGCTTTCG TAAAGCGTGC GCCATCAACG TGGCTCACAG 1000
1001 CGTATGTAGT GAAGGTATTT AGCCTCGCCG TAAATTTAAT CGCGATTGAC 1050
1051 AGTCAAGTGT TATGCGGCGC GGTCAAGTGG CTCAATCTTG AAAAGCAAAA 1100
1101 GCCAGATGGC GTATTCCAAG AGGACGCCCC AGTCATCCAC CAAGAGATGA 1150
1151 TTGGCGGCCCT CCGCAATAAC AATGAGAAGG ACATGGCGTT AACC CGT TTT 1200
1201 GTCTTAATCA GTTTACAGGA AGCCAAAGAC ATTTGTGAGG AACAGGTAAA 1250
1251 TAGTTTACCT GGGAGTATTA CGAAAGCGGG CGATTTCTTA GAAGCAAATT 1300
1301 ACATGAATCT CCAACGCTCA TACACGGTAG CGATCGCGGG ATATGCCTTA 1350
1351 GCGCAGATGG GGAGATTAAA AGGCCATTA CTGAACAAGT TTTTAACAAC 1400
1401 CGCAAAAGAC AAGAATAGGT GGGAGGACCC AGGCAAGCAA CTTTATAACG 1450
1451 TCGAAGCAAC GTCATACGCA TTATTAGCAC TCTTACAACT CAAGGACTTC 1500
1501 GACTTCGTAC CACCTGTGGT ACGGTGGCTT AACGAACAAA GGTATTACGG 1550
1551 GGGCGGATAC GGCAGCACGC AAGCGACTTT CATGGTCTTT CAAGCACTCG 1600
1601 CACAGTACCA GAAGGATGCA CCTGATCACC AAGAATTAAA CTTAGATGTC 1650
1651 AGTCTGCAGT TACCAAGTTC AGGGTCAGGT GGAGGTGGAA GTGGTGGAGG 1700
1701 TGGAAGCGGA TCTACTCCTT CAGGGAGTGG AGAACAAAAC ATGATTGGTA 1750
1751 TGACCCCTAC AGTGATCGCC GTACACTACT TAGATGAGAC AGAGCAATGG 1800
1801 GAGAAATTCG GTTTGGAGAA AAGACAGGGA GCGTTAGAAC TTATTAAAAA 1850
1851 GGGATATACA CAGCAACTCG CTTTTAGGCA GCCTAGTAGC GCATTTGCTG 1900
1901 CGTTTGTCAA AAGAGCCCCCT AGTACATGGT TAACGGCTTA CGTCGTAAAA 1950
1951 GTGTTCTCAT TAGCGGTGAA CCTGATTGCA ATCGATTTCG AGGTACTGTG 2000
2001 TGGAGCCGTG AAATGGTTAA TCTTAGAGAA ACAGAAACCT GACGGAGTGT 2050
2051 TTCAGGAAGA TGCACCTGTA ATTCACCAGG AAATGATCGG GGGCTTGAGA 2100
2101 AACAATAACG AAAAAGATAT GGCTCTGACA GCTTTCGTGC TGATTTCCCT 2150
2151 CCAAGAGGCG AAGGATATCT GCGAAGAGCA AGTGAAC TCA CTCC CAGGAT 2200
2201 CAATCACCAA GGCCGGGGAC TTTCTGGAGG CGAACTATAT GAACTTGCAG 2250
2251 AGGAGCTATA CCGTCGCAAT TGCCGGTTAC GCGCTCGCAC AAATGGGACG 2300
2301 TCTCAAAGGA CCTCTGTTAA ATAAATTCCT CACGACGGCG AAGGATAAAA 2350
2351 ACCGATGGGA AGACCCTGGG AAACAGTTGT ACAATGTAGA GGCGACCAGT 2400
2401 TATGCGCTGC TCGCGTTGCT CCAGTTGAAA GATTTTGATT TTGTCCCTCC 2450

```

2451 AGTAGTCAGA TGGTTGAATG AGCAGCGTTA CTATGGAGGG GGGTATGGAT 2500

2501 CAACACAGGC AACGTTTATG GTATTCCAGG CGTTAGCGCA ATATCAAAAA 2550

2551	GACGCGCCAG	ACCACCAGGA	GCTTAATCTC	GACGTATCAT	TACAACTCCC	2600
2601	TTCAAGCGGC	AGCGGCGGGG	GCGGGTCAGG	AGGCGGGGGT	TCTGGATCTA	2650
2651	CCCCCTCGGG	CTCCGGGGAA	CAGAACATGA	TCGGCATGAC	GCCCACGGTC	2700
2701	ATCGCTGTGC	ATTACCTGGA	TGAAACGGAG	CAGTGGGAGA	AGTTCGGCCT	2750
2751	AGAGAAGCGG	CAGGGGGCCT	TGGAGCTCAT	CAAGAAGGGG	TACACCCAGC	2800
2801	AGCTGGCCTT	CAGACAACCC	AGCTCTGCCT	TTGCGGCCTT	CGTGAAACGG	2850
2851	GCACCCAGCA	CCTGGCTGAC	CGCCTACGTG	GTCAAGGTCT	TCTCTCTGGC	2900
2901	TGTCAACCTC	ATCGCCATCG	ACTCCCAAGT	CCTCTGCGGG	GCTGTTAAAT	2950
2951	GGCTGATCCT	GGAGAAGCAG	AAGCCCAGAC	GGGTCTTCCA	GGAGGATGCG	3000
3001	CCCGTGATAC	ACCAAGAAAT	GATTGGTGGA	TTACGGAACA	ACAACGAGAA	3050
3051	AGACATGGCC	CTCACGGCCT	TTGTTCTCAT	CTCGCTGCAG	GAGGCTAAAG	3100
3101	ATATTTGCGA	GGAGCAGGTC	AACAGCCTGC	CAGGCAGCAT	CACTAAAGCA	3150
3151	GGAGACTTCC	TTGAAGCCAA	CTACATGAAC	CTACAGAGAT	CCTACACTGT	3200
3201	GGCCATTGCT	GGCTATGCTC	TGGCCAGAT	GGGCAGGCTG	AAGGGGCCTC	3250
3251	TTCTTAACAA	ATTTCTGACC	ACAGCCAAAG	ATAAGAACCG	CTGGGAGGAC	3300
3301	CCTGGTAAGC	AGCTCTACAA	CGTGGAGGCC	ACATCCTATG	CCCTCTTGGC	3350
3351	CCTACTGCAG	CTAAAAGACT	TTGACTTTGT	GCCTCCCGTC	GTGCGTTGGC	3400
3401	TCAATGAACA	GAGATACTAC	GGTGGTGGCT	ATGGCTCTAC	CCAGGCCACC	3450
3451	TTCATGGTGT	TCCAAGCCTT	GGCTCAATAC	CAAAAGGACG	CCCCTGACCA	3500
3501	CCAGGAACTG	AACCTTGATG	TGTCCCTCCA	ACTGCCCAGC	AGTGGATCTA	3550
3551	<u>ACATTGCCCA</u>	<u>ACACCAATGC</u>	<u>GTTAAGAAGC</u>	<u>AATGTCCACA</u>	<u>AAACTCCGGA</u>	3600
3601	<u>TGTTTCAGAC</u>	<u>ATCTGGACGA</u>	<u>GAGAGAAGAA</u>	<u>TGTAAGTGTC</u>	<u>TGTTGAACTA</u>	3650
3651	<u>CAAGCAGGAA</u>	<u>GGTGATAAGT</u>	<u>GTGTTGAGAA</u>	<u>CCCAAACCCT</u>	<u>ACCTGTAACG</u>	3700
3701	<u>AGAACAACGG</u>	<u>TGGATGCGAC</u>	<u>GCTGACGCTA</u>	<u>AGTGCACCGA</u>	<u>AGAAGACTCT</u>	3750
3751	<u>GGTTCTAACG</u>	<u>GAAAGAAGAT</u>	<u>TACTTGCGAA</u>	<u>TGTACTAAGC</u>	<u>CAGACTCTTA</u>	3800
3801	<u>CCCTTTGTTC</u>	<u>GATGGAATCT</u>	<u>TCTGTTCTTC</u>	<u>CTCTAACTCT</u>	<u>TCCTCTGGAT</u>	3850
3851	CC...					3852

SEQ ID 46: DNA sequence of PfMSP1.19 double cysteine mutant from example 13 (mutated codons shown in bold).

1	AGATCTAACA TTGCCCCAACA CCAATGCGTT AAGAAGCAAA TTCCACAAAA	50
51	CTCCGGATGT TTCAGACATC TGGACGAGAG AGAAGAATGG AAGTGTCTGT	100
101	TGAACTACAA GCAGGAAGGT GATAAGTGTG TTGAGAACCC AAACCCTACC	150
151	TGTAACGAGA ACAACGGTGG ATGCGACGCT GACGCTAAGT GCACCGAAGA	200
201	AGACTCTGGT TCTAACGGAA AGAAGATTAC TTGCGAATGT ACTAAGCCAG	250
251	ACTCTTACCC TTTGTTGAT GGAATCTTCT GTTCTTCCTC TAACTCTTCC	300
301	TCTGGATCC	309

SEQ ID 47: Coding sequence from pVK104-03 (mutant PfMSP1.19-human C3d₃ in PVAX3 (signal peptide in *italics*, antigen sequence underlined) from example 13).

```

706 .....ATGGA TGCAATGAAG AGAGGGCTCT GCTGTGTGCT GCTGCTGTGT 750
751 GGAGCAGTCT TCGTTTCCGC TAGATCTAAC ATTGCCCAAC ACCAATGCGT 800
801 TAAGAAGCAA ATTCCACAAA ACTCCGGATG TTTCAGACAT CTGGACGAGA 850
851 GAGAAGAATG GAAGTGTCTG TTGAACTACA AGCAGGAAGG TGATAAGTGT 900
901 GTTGAGAACC CAAACCCTAC CTGTAACGAG AACAACGGTG GATGCGACGC 950
951 TGACGCTAAG TGCACCGAAG AAGACTCTGG TTCTAACGGA AAGAAGATTA 1000
1001 CTTGCGAATG TACTAAGCCA GACTCTTACC CTTTGTTCGA TGGAATCTTC 1050
1051 TGTTCCTCCT CTAACCTCTC CTCTGGATCT ACGCCAAGCG GATCAGGCGA 1100
1101 GCAGAATATG ATCGGGATGA CACCAACCGT AATTGCGGTC CATTATCTCG 1150
1151 ACGAAACCGA ACAGTGGGAA AAATTTGGGC TCGAAAAGCG TCAAGGCGCT 1200
1201 CTCGAGTTGA TCAAGAAAGG CTACACGCAA CAGTTAGCGT TCCGTCAACC 1250
1251 ATCATCAGCG TTCGCCGCTT TCGTAAAGCG TCGCCCATCA ACGTGGCTCA 1300
1301 CAGCGTATGT AGTGAAGGTA TTTAGCCTCG CCGTAAATTT AATCGCGATT 1350
1351 GACAGTCAAG TGTTATGCGG CGCGGTCAAG TGGCTCATTC TTGAAAAGCA 1400
1401 AAAGCCAGAT GGCCTATTCC AAGAGGACGC CCCAGTCATC CACCAAGAGA 1450
1451 TGATTGGCGG CCTCCGCAAT AACAATGAGA AGGACATGGC GTTAACCGCG 1500
1501 TTTGTCTTAA TCAGTTTACA GGAAGCCAAA GACATTTGTG AGGAACAGGT 1550
1551 AAATAGTTTA CCTGGGAGTA TTACGAAAGC GGGCGATTTT TTAGAAGCAA 1600
1601 ATTACATGAA TCTCCAACGC TCATACACGG TAGCGATCGC GGGATATGCC 1650
1651 TTAGCGCAGA TGGGGAGATT AAAAGGCCCA TTAGTGAACA AGTTTTTAAC 1700
1701 AACCACAAAA GACAAGAATA GGTGGGAGGA CCCAGGCAAG CAACCTTTATA 1750
1751 ACGTCGAAGC AACGTCATAC GCATTATTAG CACTCTTACA ACTCAAGGAC 1800
1801 TTCGACTTCG TACCACCTGT GGTACGGTGG CTTAACGAAC AAAGGTATTA 1850
1851 CGGGGGCGGA TACGGCAGCA CGCAAGCGAC TTTCATGGTC TTTCAAGCAC 1900
1901 TCGCACAGTA CCAGAAGGAT GCACCTGATC ACCAAGAATT AAACCTAGAT 1950
1951 GTCAGTCTGC AGTTACCAAG TTCAGGGTCA GGTGGAGGTG GAAGTGGTGG 2000
2001 AGGTGGAAGC GGATCTACTC CTTACGGGAG TGGAGAACAA AACATGATTG 2050
2051 GTATGACCCC TACAGTGATC GCCGTACACT ACTTAGATGA GACAGAGCAA 2100
2101 TGGGAGAAAT TCGGTTTGGA GAAAAGACAG GGAGCGTTAG AACTTATTAA 2150
2151 AAAGGGATAT ACACAGCAAC TCGCTTTTAG GCAGCCTAGT AGCGCATTTG 2200
2201 CTGCGTTTGT CAAAAGAGCC CCTAGTACAT GGTTAACGGC TTACGTGCGTA 2250
2251 AAAGTGTCTT CATTAGCGGT GAACCTGATT GCAATCGATT CGCAGGTACT 2300
2301 GTGTGGAGCC GTGAAATGGT TAATCTTAGA GAAACAGAAA CCTGACGGAG 2350
2351 TGTTTCAGGA AGATGCACCT GTAATTCACC AGGAAATGAT CGGGGGCTTG 2400
2401 AGAAACAATA ACGAAAAAGA TATGGCTCTG ACAGCTTTCG TGCTGATTTT 2450

```

2451 CCTCCAAGAG GCGAAGGATA TCTGCGAAGA GCAAGTGAAC TCACTCCCAG 2500
2501 GATCAATCAC CAAGGCCGGG GACTTTCTGG AGGCGAACTA TATGAACTTG 2550

2551	CAGAGGAGCT	ATACCGTCGC	AATTGCCGGT	TACGCGCTCG	CACAAATGGG	2600
2601	ACGTCTCAAA	GGACCTCTGT	TAAATAAATT	CCTCACGACG	GCGAAGGATA	2650
2651	AAAACCGATG	GGAAGACCTT	GGGAAACAGT	TGTACAATGT	AGAGGCGACC	2700
2701	AGTTATGCGC	TGCTCGCGTT	GCTCCAGTTG	AAAGATTTTG	ATTTTGTCCC	2750
2751	TCCAGTAGTC	AGATGGTTGA	ATGAGCAGCG	TTACTATGGA	GGGGGGTATG	2800
2801	GATCAACACA	GGCAACGTTT	ATGGTATTCC	AGGCGTTAGC	GCAATATCAA	2850
2851	AAAGACGCGC	CAGACCACCA	GGAGCTTAAT	CTCGACGTAT	CATTACAAC	2900
2901	CCCTTCAAGC	GGCAGCGGCG	GGGGCGGGTC	AGGAGGCGGG	GGTTCTGGAT	2950
2951	CTACCCCTC	GGGCTCCGGG	GAACAGAACA	TGATCGGCAT	GACGCCACG	3000
3001	GTCATCGCTG	TGCATTACCT	GGATGAAACG	GAGCAGTGGG	AGAAGTTCGG	3050
3051	CCTAGAGAAG	CGGCAGGGGG	CCTTGGAGCT	CATCAAGAAG	GGGTACACCC	3100
3101	AGCAGCTGGC	CTTCAGACAA	CCCAGCTCTG	CCTTTGCGGC	CTTCGTGAAA	3150
3151	CGGGCACCCA	GCACCTGGCT	GACCGCTAC	GTGGTCAAGG	TCTTCTCTCT	3200
3201	GGCTGTCAAC	CTCATCGCCA	TCGACTCCCA	AGTCCTCTGC	GGGGCTGTTA	3250
3251	AATGGCTGAT	CCTGGAGAAG	CAGAAGCCCG	ACGGGGTCTT	CCAGGAGGAT	3300
3301	GCGCCCGTGA	TACACCAAGA	AATGATTGGT	GGATTACGGA	ACAACAACGA	3350
3351	GAAAGACATG	GCCCTCACGG	CCTTTGTTCT	CATCTCGCTG	CAGGAGGCTA	3400
3401	AAGATATTTG	CGAGGAGCAG	GTCAACAGCC	TGCCAGGCAG	CATCACTAAA	3450
3451	GCAGGAGACT	TCCTTGAAGC	CAACTACATG	AACCTACAGA	GATCCTACAC	3500
3501	TGTGGCCATT	GCTGGCTATG	CTCTGGCCCA	GATGGGCAGG	CTGAAGGGGC	3550
3551	CTCTTCTTAA	CAAATTTCTG	ACCACAGCCA	AAGATAAGAA	CCGCTGGGAG	3600
3601	GACCCTGGTA	AGCAGCTCTA	CAACGTGGAG	GCCACATCCT	ATGCCCTCTT	3650
3651	GGCCCTACTG	CAGCTAAAAG	ACTTTGACTT	TGTGCCTCCC	GTCGTGCGTT	3700
3701	GGCTCAATGA	ACAGAGATAC	TACGGTGGTG	GCTATGGCTC	TACCCAGGCC	3750
3751	ACCTTCATGG	TGTTCCAAGC	CTTGGCTCAA	TACCAAAAGG	ACGCCCTTGA	3800
3801	CCACCAGGAA	CTGAACCTTG	ATGTGTCCCT	CCAAGTCCCC	AGCAGTGGAT	3850
3851	CC...					3852

SEQ ID 48: Coding sequence from pVK104-04 (human C3d₃-mutant PfMSP1.19 in PVAX3 (signal peptide in *italics*, antigen sequence underlined) from example 13).

```

706   . . . . .ATGGA TGCAATGAAG AGAGGGCTCT GCTGTGTGCT GCTGCTGTGT      750
751   GGAGCAGTCT TCGTTTCCGC TAGATCTACG CCAAGCGGAT CAGGCGAGCA      800
801   GAATATGATC GGGATGACAC CAACCGTAAT TGCGGTCCAT TATCTCGACG      850
851   AAACCGAACA GTGGGAAAAA TTTGGGCTCG AAAAGCGTCA AGGCGCTCTC      900
901   GAGTTGATCA AGAAAAGGCTA CACGCAACAG TTAGCGTTCC GTCAACCATC      950
951   ATCAGCGTTC GCCGCTTTTCG TAAAGCGTGC GCCATCAACG TGGCTCACAG     1000
1001  CGTATGTAGT GAAGGTATTT AGCCTCGCCG TAAATTTAAT CGCGATTGAC     1050
1051  AGTCAAGTGT TATGCGGCGC GGTCAAGTGG CTCATTCTTG AAAAGCAAAA     1100
1101  GCCAGATGGC GTATTCCAAG AGGACGCCCC AGTCATCCAC CAAGAGATGA     1150
1151  TTGGCGGCCT CCGCAATAAC AATGAGAAGG ACATGGCGTT AACCGCGTTT     1200
1201  GTCTTAATCA GTTTACAGGA AGCCAAAGAC ATTTGTGAGG AACAGGTAAA     1250
1251  TAGTTTACCT GGGAGTATTA CGAAAGCGGG CGATTTCTTA GAAGCAAATT     1300
1301  ACATGAATCT CCAACGCTCA TACACGGTAG CGATCGCGGG ATATGCCTTA     1350
1351  GCGCAGATGG GGAGATTAAA AGGCCCATTA CTGAACAAAGT TTTTAACAAC     1400
1401  CGCAAAAGAC AAGAATAGGT GGGAGGACCC AGGCAAGCAA CTTTATAACG     1450
1451  TCGAAGCAAC GTCATACGCA TTATTAGCAC TCTTACAACCT CAAGGACTTC     1500
1501  GACTTCGTAC CACCTGTGGT ACGGTGGCTT AACGAACAAA GGTATTACGG     1550
1551  GGGCGGATAC GGCAGCACGC AAGCGACTTT CATGGTCTTT CAAGCACTCG     1600
1601  CACAGTACCA GAAGGATGCA CCTGATCACC AAGAATTAAA CTTAGATGTC     1650
1651  AGTCTGCAGT TACCAAGTTC AGGGTCAGGT GGAGGTGGAA GTGGTGGAGG     1700
1701  TGGAAGCGGA TCTACTCCTT CAGGGAGTGG AGAACAAAAC ATGATTGGTA     1750
1751  TGACCCCTAC AGTGATCGCC GTACACTACT TAGATGAGAC AGAGCAATGG     1800
1801  GAGAAATTCT GTTTGGAGAA AAGACAGGGA GCGTTAGAAC TTATTAAAAA     1850
1851  GGGATATACA CAGCAACTCG CTTTTAGGCA GCCTAGTAGC GCATTTGCTG     1900
1901  CGTTTGTCAA AAGAGCCCTT AGTACATGGT TAACGGCTTA CGTCGTAAAA     1950
1951  GTGTTCTCAT TAGCGGTGAA CCTGATTGCA ATCGATTTCG AGGTACTGTG     2000
2001  TGGAGCCGTG AAATGGTTAA TCTTAGAGAA ACAGAAACCT GACGGAGTGT     2050
2051  TTCAGGAAGA TGCACCTGTA ATTCACCAGG AAATGATCGG GGGCTTGAGA     2100
2101  AACATAACG AAAAAAGATAT GGCTCTGACA GCTTTCGTGC TGATTTCCCT     2150
2151  CCAAGAGGCG AAGGATATCT GCGAAGAGCA AGTGAACCTA CTCCCAGGAT     2200
2201  CAATCACCAA GGCCGGGGAC TTTCTGGAGG CGAACTATAT GAACTTGCAG     2250
2251  AGGAGCTATA CCGTCGCAAT TGCCGGTTAC GCGCTCGCAC AAATGGGACG     2300
2301  TCTCAAAGGA CCTCTGTTAA ATAAATTCCT CACGACGGCG AAGGATAAAA     2350
2351  ACCGATGGGA AGACCCTGGG AACAGTTGT ACAATGTAGA GGCGACCAGT     2400
2401  TATGCGCTGC TCGCGTTGCT CCAGTTGAAA GATTTTGATT TTGTCCCTCC     2450
2451  AGTAGTCAGA TGGTTGAATG AGCAGCGTTA CTATGGAGGG GGGTATGGAT     2500

```

2501 CAACACAGGC AACGTTTATG GTATTCCAGG CGTTAGCGCA ATATCAAAAA 2550

2551	GACGCGCCAG	ACCACCAGGA	GCTTAATCTC	GACGTATCAT	TACAAC TCCC	2600
2601	TTCAAGCGGC	AGCGGCGGGG	GCGGGTCAGG	AGGCGGGGGT	TCTGGATCTA	2650
2651	CCCCCTCGGG	CTCCGGGGAA	CAGAACATGA	TCGGCATGAC	GCCCACGGTC	2700
2701	ATCGCTGTGC	ATTACCTGGA	TGAAACGGAG	CAGTGGGAGA	AGTTCGGCCT	2750
2751	AGAGAAGCGG	CAGGGGGCCT	TGGAGCTCAT	CAAGAAGGGG	TACACCCAGC	2800
2801	AGCTGGCCTT	CAGACAACCC	AGCTCTGCCT	TTGCGGCCTT	CGTGAAACGG	2850
2851	GCACCCAGCA	CCTGGCTGAC	CGCCTACGTG	GTCAAGGTCT	TCTCTCTGGC	2900
2901	TGTCAACCTC	ATCGCCATCG	ACTCCCAAGT	CCTCTGCGGG	GCTGT TAAAT	2950
2951	GGCTGATCCT	GGAGAAGCAG	AAGCCCGACG	GGGTCTTCCA	GGAGGATGCG	3000
3001	CCCGTGATAC	ACCAAGAAAT	GATTGGTGGA	TTACGGAACA	ACAACGAGAA	3050
3051	AGACATGGCC	CTCACGGCCT	TTGTTCTCAT	CTCGCTGCAG	GAGGCTAAAG	3100
3101	ATATTTGCGA	GGAGCAGGTC	AACAGCCTGC	CAGGCAGCAT	CACTAAAGCA	3150
3151	GGAGACTTCC	TTGAAGCCAA	CTACATGAAC	CTACAGAGAT	CCTACACTGT	3200
3201	GGCCATTGCT	GGCTATGCTC	TGGCCCAGAT	GGGCAGGCTG	AAGGGGCCTC	3250
3251	TTCTTAACAA	ATTTCTGACC	ACAGCCAAAG	ATAAGAACCG	CTGGGAGGAC	3300
3301	CCTGGTAAGC	AGCTCTACAA	CGTGGAGGCC	ACATCCTATG	CCCTCTTGGC	3350
3351	CCTACTGCAG	CTAAAAGACT	TTGACTTTGT	GCCTCCCGTC	GTGCGTTGGC	3400
3401	TCAATGAACA	GAGATACTAC	GGTGGTGGCT	ATGGCTCTAC	CCAGGCCACC	3450
3451	TTCATGGTGT	TCCAAGCCTT	GGCTCAATAC	CAAAAGGACG	CCCCTGACCA	3500
3501	CCAGGAAC TG	AACCTTGATG	TGTCCCTCCA	ACTGCCCAGC	AGTGGATCTA	3550
3551	<u>ACATTGCCCA</u>	<u>ACACCAATGC</u>	<u>GTTAAGAAGC</u>	<u>AAATTCCACA</u>	<u>AAACTCCGGA</u>	3600
3601	<u>TGTTTCAGAC</u>	<u>ATCTGGACGA</u>	<u>GAGAGAAGAA</u>	<u>TGGAAGTGTC</u>	<u>TGTTGAACTA</u>	3650
3651	<u>CAAGCAGGAA</u>	<u>GGTGATAAGT</u>	<u>GTGTTGAGAA</u>	<u>CCCAAACCCT</u>	<u>ACCTGTAACG</u>	3700
3701	<u>AGAACAACGG</u>	<u>TGGATGCGAC</u>	<u>GCTGACGCTA</u>	<u>AGTGCACCGA</u>	<u>AGAAGACTCT</u>	3750
3751	<u>GGTTCTAACG</u>	<u>GAAAGAAGAT</u>	<u>TACTTGCGAA</u>	<u>TGTACTAAGC</u>	<u>CAGACTCTTA</u>	3800
3801	<u>CCCTTTGTTC</u>	<u>GATGGAATCT</u>	<u>TCTGTTCTTC</u>	<u>CTCTAACTCT</u>	<u>TCCTCTGGAT</u>	3850
3851	CC...					3852

SEQ ID 49:

DNA sequence of third human C3d variant (Example 14)

1	ACACCGTCTG	GTAGCGGTGA	GCAAAATATG	ATAGGAATGA	CTCCGACTGT	50
51	TATAGCAGTT	CACTATTTAG	ACGAGACTGA	ACAATGGGAA	AAGTTTGGAC	100
101	TGGAAAAAAG	GCAAGGTGCA	CTGGAATTAA	TAAAAAAGG	TTATACGCAG	150
151	CAACTAGCGT	TCAGGCAGCC	GTCGAGCGCT	TTCGCAGCAT	TTGTCAAGAG	200
201	GGCTCCGTCC	ACTTGGTTGA	CGGCATATGT	CGTGAAAGTT	TTTAGTTTGG	250
251	CAGTTAACTT	GATAGCGATC	GATAGCCAGG	TTTTGTGTGG	TGCAGTAAAG	300
301	TGGTTGATAC	TCGAAAAGCA	AAAGCCGGAT	GGTGTTTTTC	AAGAAGACGC	350
351	CCCGGTTATC	CATCAGGAGA	TGATCGGAGG	TCTGAGGAAT	AATAATGAAA	400
401	AGGATATGGC	ATTGACTGCA	TTCGTATTGA	TAAGCTTGCA	AGAAGCAAAG	450
451	GACATATGTG	AAGAACAAGT	TAATTCCTTG	CCGGGTTCCT	TAACAAAGGC	500
501	TGGTGATTTT	CTCGAGGCTA	ATTATATGAA	TCTGCAACGA	AGTTATACAG	550
551	TTGCTATAGC	AGGGTACGCA	CTCGCTCAAA	TGGGTCGCTT	GAAGGGTCCG	600
601	CTCCTGAATA	AGTTCTTGAC	TACTGCTAAG	GACAAAATA	GATGGGAAGA	650
651	TCCGGGAAAA	CAACTGTATA	ATGTTGAAGC	TACTAGCTAC	GCTTTGCTGG	700
701	CTCTGTTGCA	ACTGAAGGAT	TTCGATTTTC	TTCCCCCGGT	TGTTAGGTGG	750
751	TTAAACGAGC	AACGCTATTA	TGGCGGAGGT	TACGGGTCGA	CTCAAGCTAC	800
801	ATTTATGGTT	TTTCAGGCTC	TGGCCCAGTA	TCAGAAAGAT	GCTCCCGATC	850
851	ATCAAGAGCT	CAATCTGGAC	GTTAGCTTGC	AGTTGCCG		988

SEQ ID 50: Nucleotide sequence of PCR primer FARM 1 (example 16)

1 TGYGGRGARC AGAACATGAT YGGCATG 27

SEQ ID 51: Nucleotide sequence of PCR primer FARM 2 (example 16)

1 CCGTAGTATC TYASNTCRTT GAGCCA 26

SEQ ID 52: Nucleotide sequence of PCR primer FARM 3 (example 16)

1 GGAGTCTTCG AGGAGAATGG GCC 23

SEQ ID 53: Nucleotide sequence of PCR primer FARM 4 (example 16)

1 GTGTGTCWGG RRCRAAGCCR GTCATCAT 28

SEQ ID 54: Nucleotide sequence of PCR primer FARM 5 (example 16)

1 GTRATGCAGG ACTTCTTCAT YGACCTG 27

SEQ ID 55: Nucleotide sequence of PCR primer FARM 6 (example 16)

1 GGCTGTCAGG GACACGTCCTT TCTC 24

SEQ ID 56: Nucleotide sequence of PCR primer FARM 7 (example 16)

1 GCAAGGGACC CCMGTGGCCC AGATG 25

SEQ ID 57: Nucleotide sequence of PCR primer FARM 8 (example 16)

1 GYCACCACCG ACAAKGTGCC TTG 23

(for SEQ ID 50- SEQ ID 57 R=G/A, Y=C/T, W=A/T, S=G/C, K=G/T, M=A/C
N=A/C/G/T)

SEQ ID 58:

DNA sequence of wild-type rhesus macaque C3d (Example 16)

1	ACCCCTCGG	GCTGCGGAGA	ACAGAACATG	ATCACCATGA	CGCCCACAGT	50
51	CATCGCTGTG	CATTACCTGG	ATGAAACGGA	ACAGTGGGAG	AAGTTCGGCC	100
101	CGGAGAAGCG	GCAGGGGGCC	TTGGAGCTCA	TCAAGAAGGG	GTACACCCAG	150
151	CAGCTGGCCT	TCAGACAACC	CAGCTCTGCC	TTTGGCGCCT	TCCTGAACCG	200
201	GGCACCAGC	ACCTGGCTGA	CCGCCTACGT	GGTCAAGGTC	TTCTCTCTGG	250
251	CTGTCAACCT	CATTGCCATC	GACTCCCAGG	TCCTCTGCGG	GGCTGTTAAA	300
301	TGGCTGATCC	TGGAGAAGCA	GAAGCCCGAC	GGGGTCTTCC	AGGAGGATGC	350
351	GCCCCGTGATA	CATCAAGAAA	TGACTGGTGG	ATTCCGGAAC	ACCAACGAGA	400
401	AAGACATGGC	CCTCACGGCC	TTTGTTCTCA	TCTCGCTGCA	AGAGGCTAAA	450
451	GAGATTTGCG	AGGAGCAGGT	CAACAGCCTG	CCCGGCAGCA	TCTACTAAAGC	500
501	AGGAGACTTC	CTTGAAGCCA	ACTACATGAA	CCTACAGAGA	TCCTACACTG	550
551	TGGCCATCGC	TGCCTATGCC	CTGGCCCAGA	TGGGCAGGCT	GAAGGGACCT	600
601	CTTCTCAACA	AATTTCTGAC	CACAGCCAAA	GATAAGAACC	GCTGGGAGGA	650
651	GCCTGGTCAG	CAGCTCTACA	ATGTGGAGGC	CACATCCTAT	GCCCTCTTGG	700
701	CCCTACTGCA	GCTAAAAGAC	TTTGACTTTG	TGCCTCCCGT	CGTGCGTTGG	750
751	CTCAATGAAC	AGAGATACTA	CGGTGGTGGC	TATGGCTCTA	CCCAGGCCAC	800
801	CTTCATGGTG	TTCCAAGCCT	TGGCTCAATA	CCAAAAGGAT	GTCCCTGATC	850
851	ACAAGGAACT	GAACCTGGAT	GTGTCCCTCC	AACTGCCC		888

SEQ ID 59:

DNA sequence of first variant of rhesus macaque C3d (Example 17)

1	ACGCCAAGCG	GATCAGGCGA	GCAGAATATG	ATCACTATGA	CACCAACCGT	50
51	AATTGCGGTC	CATTATCTCG	ACGAAACCGA	ACAGTGGGAA	AAATTTGGGC	100
101	CGGAAAAGCG	TCAAGGCGCT	CTCGAGTTGA	TCAAGAAAGG	CTACACGCAA	150
151	CAGTTAGCGT	TCCGTCAACC	ATCATCAGCG	TTCGCCGCTT	TCCTGAATCG	200
201	TGCGCCATCA	ACGTGGCTCA	CAGCGTATGT	AGTGAAGGTA	TTTAGCCTCG	250
251	CCGTAAATTT	AATCGCGATT	GACAGTCAAG	TGTTATGCGG	CGCGGTCAAG	300
301	TGGCTCATTC	TTGAAAAGCA	AAAGCCAGAT	GGCGTATTCC	AAGAGGACGC	350
351	CCCAGTCATC	CACCAAGAGA	TGACAGGCGG	CTTTCGCAAT	ACTAATGAGA	400
401	AGGACATGGC	GTTAACCGCG	TTTGTCTTAA	TCAGTTTACA	GGAAGCCAAA	450
451	GAAATTTGTG	AGGAACAGGT	AAATAGTTTA	CCTGGGAGTA	TTACGAAAGC	500
501	GGGCGATTTT	TTAGAAGCAA	ATTACATGAA	TCTCCAACGC	TCATACACGG	550
551	TAGCGATCGC	GGCTTATGCC	TTAGCGCAGA	TGGGGAGATT	AAAAGGCCCA	600
601	TTACTGAACA	AGTTTTTAAC	AACCGCAAAA	GACAAGAATA	GGTGGGAGGA	650
651	ACCAGGCCAA	CAACTTTATA	ACGTCGAAGC	AACGTCATAC	GCATTATTAG	700
701	CACTCTTACA	ACTCAAGGAC	TTCGACTTCG	TACCACCTGT	GGTACGGTGG	750
751	CTTAACGAAC	AAAGGTATTA	CGGGGGCGGA	TACGGCAGCA	CGCAAGCGAC	800
801	TTTCATGGTC	TTTCAAGCAC	TCGCACAGTA	CCAGAAGGAT	GTTCTTGATC	850
851	ACAAGGAATT	AAACTTAGAT	GTCAGTCTGC	AGTTACCA		888

SEQ ID 60:

DNA sequence of second variant of rhesus macaque C3d (Example 17)

1	ACTCCTTCAG	GGAGTGGAGA	ACAAAACATG	ATTACAATGA	CCCCCTACAGT	50
51	GATCGCCGTA	CACTACTTAG	ATGAGACAGA	GCAATGGGAG	AAATTCGGTC	100
101	CCGAGAAAAG	ACAGGGAGCG	TTAGAACTTA	TTAAAAAGGG	ATATACACAG	150
151	CAACTCGCTT	TTAGGCAGCC	TAGTAGCGCA	TTTGCTGCGT	TTCTCAACAG	200
201	AGCCCCTAGT	ACATGGTTAA	CGGCTTACGT	CGTAAAAGTG	TTCTCATTAG	250
251	CGGTGAACCT	GATTGCAATC	GATTGCGAGG	TACTGTGTGG	AGCCGTGAAA	300
301	TGGTTAATCT	TAGAGAAACA	GAAACCTGAC	GGAGTGTTC	AGGAAGATGC	350
351	ACCTGTAATT	CACCAGGAAA	TGACCGGGGG	CTTCAGAAAC	ACAAACGAAA	400
401	AAGATATGGC	TCTGACAGCT	TTCGTGCTGA	TTTCCCTCCA	AGAGGCGAAG	450
451	GAGATCTGCG	AAGAGCAAGT	GAACTCACTC	CCAGGATCAA	TCACCAAGGC	500
501	CGGGGACTTT	CTGGAGGCGA	ACTATATGAA	CTTGCGAGAG	AGCTATACCG	550
551	TCGCAATTGC	CGCATACGCG	CTCGCACAAA	TGGGACGTCT	CAAAGGACCT	600
601	CTGTTAAATA	AATTCCTCAC	GACGGCGAAG	GATAAAAACC	GATGGGAAGA	650
651	ACCTGGGCAA	CAGTTGTACA	ATGTAGAGGC	GACCAGTTAT	GCGCTGCTCG	700
701	CGTTGCTCCA	GTTGAAAGAT	TTTGATTTTG	TCCCTCCAGT	AGTCAGATGG	750
751	TTGAATGAGC	AGCGTTACTA	TGGAGGGGGG	TATGGATCAA	CACAGGCAAC	800
801	GTTTATGGTA	TTCCAGGCGT	TAGCGCAATA	TCAAAAAGAC	GTGCCAGACC	850
851	ACAAAGAGCT	TAATCTCGAC	GTATCATTAC	AACTCCCT		888

SEQ ID 61:

DNA sequence of third variant of rhesus macaque C3d (Example 17)

1	ACACCGTCTG	GTAGCGGTGA	GCAAAATATG	ATAACCATGA	CTCCGACTGT	50
51	TATAGCAGTT	CACTATTTAG	ACGAGACTGA	ACAATGGGAA	AAGTTTGGAC	100
101	CGGAAAAAAG	GCAAGGTGCA	CTGGAATTAA	TAAAAAAAGG	TTATACGCAG	150
151	CAACTAGCGT	TCAGGCAGCC	GTCCAGCGCT	TTCGCAGCAT	TTCTGAACAG	200
201	GGCTCCGTCC	ACTTGGTTGA	CGGCATATGT	CGTGAAAGTT	TTTAGTTTGG	250
251	CAGTTAACTT	GATAGCGATC	GATAGCCAGG	TTTGTGTGG	TGCAGTAAAG	300
301	TGGTTGATAC	TCGAAAAGCA	AAAGCCGGAT	GGTGTTTTTC	AAGAAGACGC	350
351	CCCGGTTATC	CATCAGGAGA	TGACTGGAGG	TTTCAGGAAT	ACCAATGAAA	400
401	AGGATATGGC	ATTGACTGCA	TTCGTATTGA	TAAGCTTGCA	AGAAGCAAAG	450
451	GAGATATGTG	AAGAACAAGT	TAATTCCCTG	CCGGGTTCCT	TAACAAAGGC	500
501	TGGTGATTTT	CTCGAGGCTA	ATTATATGAA	TCTGCAACGA	AGTTATACAG	550
551	TTGCTATAGC	AGCCTACGCA	CTCGCTCAAA	TGGGTCGCTT	GAAGGGTCCG	600
601	CTCCTGAATA	AGTTCTTGAC	TACTGCTAAG	GACAAAAATA	GATGGGAAGA	650
651	GCCGGGACAG	CAACTGTATA	ATGTTGAAGC	TACTAGCTAC	GCTTTGCTGG	700
701	CTCTGTTGCA	ACTGAAGGAT	TTGATTTTCG	TTCCCCCGGT	TGTTAGGTGG	750
751	TTAAACGAGC	AACGCTATTA	TGGCGGAGGT	TACGGGTCGA	CTCAAGCTAC	800
801	ATTTATGGTT	TTTCAGGCTC	TGGCCCAGTA	TCAGAAAGAT	GTCCCCGATC	850
851	ATAAGGAGCT	CAATCTGGAC	GTTAGCTTGC	AGTTGCCG		888

INTERNATIONAL SEARCH REPORT

Inter | Application No

PCT/GB 01/01599

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C12N15/62 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, STRAND, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 35260 A (ROWLING PAMELA JANE ELIZABETH ;ADPROTECH PLC (GB); COX VIVIENNE FR) 15 July 1999 (1999-07-15)	1-26
Y	pages 2,3,4,14,15,47; page 14, line 20-21, page 35, last line; examples and claims	21-23
Y	WO 96 17625 A (DEMPSEY PAUL WALTER ;FEARON DOUGLAS THOMAS (GB); LYNXVALE LTD (GB)) 13 June 1996 (1996-06-13) abstract; page 2,4,5,7,25; claims	21-23
Y	WO 99 07839 A (VLAAMS INTERUNIV INST BIOTECH ;MIN JOU WILLY (BE); FIERIS WALTER (B) 18 February 1999 (1999-02-18) page 7, 2nd paragraph, last paragraph; page 20, line 34,35; pages 49,52,53,54,57; Fig. 17,18	21-23
	-/-	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

G document member of the same patent family

Date of the actual completion of the international search

27 August 2001

Date of mailing of the international search report

06/09/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer:

Holtorf, S

INTERNATIONAL SEARCH REPORT

Interr I Application No
PCT/GB 01/01599

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 88 03533 A (SYNTRO CORP) 19 May 1988 (1988-05-19) page 2, line 31-36; page 5; page 10, line 30-33; page 13, line 24-31; example 2; especially example 3 /page 45 ---	
A	WO 93 12257 A (HYBRITECH INC) 24 June 1993 (1993-06-24) abstract; page 8 ---	
A	DÜBEL, S., ET AL.: "bifunctional and multimeric complexes of streptavidin fused to single chain antibodies (scFv)" JOURNAL OF IMMUNOLOGICAL METHODS, vol. 178, 1995, pages 201-209, XP002102911 abstract; figure 1 ---	
A	MORGENSTERN J P ET AL: "ADVANCED MAMMALIAN GENE TRANSFER: HIGH TITRE RETROVIRAL VECTORS WITH MULTIPLE DRUG SELECTION MARKERS AND A COMPLEMENTARY HELPER -FREE PACKAGING CELL LINE" NUCLEIC ACIDS RESEARCH, vol. 18, no. 12, 1 January 1990 (1990-01-01), pages 3587-3596, XP000652091 page 3593, right-hand column -page 3594, left-hand column ---	
A	DE BRUIJN M H L ET AL: "HUMAN COMPLEMENT C-3 COMPLEMENTARY DNA CODING SEQUENCE AND DERIVED PRIMARY STRUCTURE" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 82, no. 3, 1985, pages 708-712, XP002175583 1985 ISSN: 0027-8424 -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

Interr I Application No
PCT/GB 01/01599

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9935260 A	15-07-1999	AU 1776299 A EP 1042467 A	26-07-1999 11-10-2000
WO 9617625 A	13-06-1996	AU 709990 B AU 4120496 A CA 2207000 A EP 0797452 A JP 10509723 T US 6238670 B	09-09-1999 26-06-1996 13-06-1996 01-10-1997 22-09-1998 29-05-2001
WO 9907839 A	18-02-1999	AU 9341698 A EP 0996717 A	01-03-1999 03-05-2000
WO 8803533 A	19-05-1988	AU 612340 B AU 1052888 A DK 369788 A EP 0293443 A FI 883082 A JP 10014586 A JP 2000135092 A JP 1501755 T NO 882946 A NZ 222450 A US 5723588 A US 5243038 A US 5641648 A US 6140072 A US 5770697 A US 6184348 B US 6018030 A US 5514581 A US 5773249 A US 5830713 A	11-07-1991 01-06-1988 01-07-1988 07-12-1988 28-06-1988 20-01-1998 16-05-2000 22-06-1989 31-08-1988 26-09-1990 03-03-1998 07-09-1993 24-06-1997 31-10-2000 23-06-1998 06-02-2001 25-01-2000 07-05-1996 30-06-1998 03-11-1998
WO 9312257 A	24-06-1993	AU 3274793 A US 5512463 A	19-07-1993 30-04-1996

THIS PAGE BLANK (USPTO)